Logistics Management Institute

Executive Summary

In order to effectively meet the intent of the Toxic Substances Control Act (TSCA) of 1977 and serve the public interest, EPA's Office of Pollution Prevention and Toxics (OPPT) must modernize and streamline the receipt, processing, and distribution of this information. Added data collection initiatives, primarily Chemical Right to Know, will mean that OPPT's over-taxed information infrastructure will suffer additional inefficiencies. Without a serious reengineering effort, OPPT will have trouble meeting its statutory requirements.

Reengineering of TSCA processes is enabled by electronically capturing the data as far "upstream" in the submission process as possible. This means getting the manufacturer or importer to submit the data electronically. Toward this end, OPPT must establish user-friendly electronic reporting methods such as direct submission of data and documents over the Internet. This includes the use of simple HTML forms for direct data entry as well as Portable Document Format (PDF) for document submission and storage.

In addition, OPPT must establish a "central registration" authority consistent with efforts by the Office of Environmental Information (OEI). This authority will provide the framework for authenticating submissions and establishing non-repudiation for electronically submitted documents.

Once the data are captured electronically, further process improvement is possible. This includes the following recommendations which are discussed in Chapter 3:

◆ Reengineer Confidential Business Information Center (CBIC) processing. Electronic receipt into a reliable application obviates the requirement for some CBIC processing including data entry into Confidential Business Information Tracking System (CBITS) and hardcopy workflow. In turn this reduces costs, eliminates delays, and provides a more secure architecture.

◆ Consolidate and connect core TSCA databases. All processes of the core TSCA program are related enough to warrant a single connected data

store and significant efficiencies can be gained from the elimination of the fragmented, redundant architecture that currently exists. Existing fragmented internal databases should be consolidated into a connected core TSCA data store. To expand the reach and utility of TSCA information, the core TSCA data store should incorporate international information gathered under the Chemical Right-to-Know initiative and the Screening Information Data Set (SIDS) gathered for the Organization for Economic Cooperation and Development (OECD).

◆ Work from the connected core TSCA data store. Existing workflow applications can quickly and easily be modified to run off a single connected data store and new workflow applications can be generated to meet processing needs as necessary. They would draw source data from the core TSCA data store which better meets the needs of OPPT. The workflow applications would allow faster, cheaper processing of the information, and give management much better visibility of the processes.

Underlying all these recommendations is the need to reevaluate security requirements in an environment where much of today's business is conducted over the web. While TSCA data can be very sensitive, the current networking architecture is a hindrance to process improvement. The fact is that with current technology OPPT can improve on current security in an architecture that includes submission over the Internet.

Before these changes can be realized, however, two organizational and cultural issues must be addressed. The first of these is a commitment from management. Based on our interviews, we discovered that OPPT personnel were universal in recognizing the need for improvement and welcome its arrival, but also need to be convinced that a specific change is the "right direction." The commitment must come from management and must be demonstrated by follow-through. Based on interviews with OPPT personnel, the follow-through is what has been lacking on previous projects.

The second issue that must be resolved is the lack of a clear line of responsibility for section 4, 5, and 8 processing. There must be a clear answer to the questions "who's responsible for" or "who coordinates" these efforts. OPPT must assign both the coordinating and data ownership role for each type of submission. Without an individual or single organization "in charge," managing the existing process is difficult and effectively implementing process improvement is even more difficult. By establishing simple electronic reporting methods, and by improving internal data storage and processing through data centralization, OPPT can be well positioned to fill its role as the TSCA information broker.

Chapter 1

Reengineering TSCA Processing – An Overview

BACKGROUND

The Government Performance and Results Act of 1993 (GPRA) requires federal agencies to focus on their mission and goals and how to achieve them more effectively and efficiently. Implementing GPRA requires agencies to set goals, measure performance, and report on accomplishments, all with the intent of producing results. We believe the Toxic Substances Control Act (TSCA) of 1977 is a good example of a process, which can benefit from this methodology.

TSCA results in the collection, analysis, and distribution to the public of large amounts of industry data. The Office of Pollution Prevention and Toxics (OPPT) has succeeded in meeting the basic intent of TSCA for many years, primarily by managing the data flow through incremental, evolutionary improvements in information management. Over time, however, this approach has led to an information processing architecture characterized by a proliferation of stand-alone systems and stovepipe business processes. Advances in technologies and data management approaches have largely passed the process by to the point where TSCA is almost completely dependent on paper and the costs associated with manual processing. More importantly, it unnecessarily delays the distribution of vital health and safety information. With new reporting initiatives, such as Chemical Right to Know (CRTK), a system already bending under the strain of current volume may simply break.

This report is based on a recognized need within OPPT to examine TSCA processing with the intent of improving, or, if warranted, overhauling the entire process. This reengineering effort began in August of 1997. Since that time LMI has reviewed existing literature and conducted dozens of interviews with EPA personnel and individuals from industry. This report provides our conclusions and recommendations for reengineering TSCA processing.

SCOPE

The study focuses on those elements of core TSCA that represent the bulk of processing effort: section 5 (the Premanufacture Notice), sections 4 and 8 [Health and Safety, High Production Volume (HPV) and For-Your-Information (FYI) reports], and section 12b (the Export Notice). It provides recommendations for improved processing with emphasis on better information management and architecture. While we provide general recommendations for workflow and assign general responsibility at the branch level, it was not our intent to identify day-to-day responsibility of the subject matter experts (SME). This report does not make recommendations by branch at the data element level, nor does it provide functional requirements for system development and implementation,

which is the next step in reengineering the TSCA system.

Problems with Current TSCA Processing

The principal findings of this study demonstrate that TSCA processing is too dependent upon paper and antiquated data storage methods. The processing architecture is not based on a comprehensive analysis of processing requirements but instead evolved over time to meet short term requirements with the existing technology. Findings gathered from interviews leads to conclusions addressed in the following paragraphs.

Program Management

Gathering data for this study represented our biggest challenge. Very few, if any, of the EPA personnel we spoke with could provide us with a clear picture even at a macro level of processing under TSCA sections 4, 5, or 8. With the exception of section 12b, processing is cross-functional within OPPT. Sections 4, 5, and 8 require analysis and expertise from a number of different divisions and branches within those divisions. It is not apparent, however, if there is an individual or branch with overall responsibility to manage or coordinate other branches' efforts. Evidence of this comes from asking several EPA personnel, for example, "who's in charge of section 8d processing" or "who owns health and safety data." The answers generally did not provide a clear picture of authority and responsibility within OPPT.

While this study is primarily interested in data management, this may be the most significant finding. Without an individual or group clearly assigned coordination responsibility, it is difficult to effectively manage an existing process, and even more difficult to enact process change. Before there can be improvement in data management, there must be a definitive answer to the question "who owns the data?" The lack of clear program management responsibility contributes to the findings detailed in the next several sections.

Processing is Paper Intensive

While the possibility of accepting electronic TSCA submissions is being explored, all submissions are currently only accepted as paper. The proliferation of copies begins with the submitter who, depending upon the specific submission, must provide multiple copies of the same submission, including a "sanitized" copy with confidential business information (CBI) blacked out. Upon receipt at the Confidential Business Information Center (CBIC), up to seven more copies of the original submission may be made and propagated throughout the office. The result is processes that are costly and slow.

Processing is Slow and Labor Intensive

The value of information changes over time, and in many cases, such as with

TSCA, it may degrade over time. Because submissions are received and routed as paper, the processes are predictably slow and labor intensive. In the case of the Export Notice, for example, the time from submission to receipt to the point where a letter is mailed to the country of import often takes longer than the 5 days allowed by law, thereby lessening the value of the information to that country. In the case of Health and Safety studies, the contractor performing data entry into the tracking system at one time had a two to four year backlog of studies awaiting processing.

Data Storage is Fragmented and Redundant

Because of the existing architecture's evolutionary development, data storage across all three processes is exceptionally fragmented, most notably in the case of the Premanufacture Notice (PMN). Currently PMN data are entered into dozens of stand-alone systems and models, but no single system or combination of systems can electronically reproduce the PMN! At the Non Confidential Information Center (NCIC), docket management is also characterized by a lack of integration. In this case, the stand-alone database used to track hard copy information about each TSCA rule creates a separate dBase file for the more than 100 dockets managed by the NCIC. This means that the NCIC would have to open more than 100 files in order to access information organized in any way other than by docket (by chemical, for example).

Fragmented data storage leads to inefficiencies such as version conflict and redundant data entry. Perhaps more importantly, it makes it difficult for management to extract meaningful information from the data. Additional inefficiencies include the need to maintain and update a variety of different systems across different platforms.

Public Access is Not Optimal

The future of OPPT is acting as an industry to public information broker for TSCA information. Unfortunately, much of the information that is currently distributed is not as timely as it could be, and the methods used for distribution do not favor universal access. The lack of timeliness is primarily a function of processing delays as previously discussed. The public's primary access method to TSCA information, however, is through the public docket at the NCIC. Requests must be submitted in writing, and, while the staff manages to process them expeditiously, the requests must be juggled with the task of filing and indexing of paper records. Copies must then be made and mailed to the requesting organization.

Document Tracking and CBI Take Significant Resources

A significant amount of the effort required to process all TSCA submissions is

directed toward activities not directly related to evaluating a chemical's potential health and environmental effects. Section 14 of TSCA allows a submitter to claim much of the information CBI and thereby imposes the burden of strict document management on OPPT. This necessitates the actions of the Confidential Business Information Center (CBIC) including data entry into the CBI Tracking System (CBITS). The effort required to sort, copy, rout, and perform data entry is significant and contributes to delays in processing of anywhere from two to five days. In the case of the Export Notice, where 75% of submissions are non-CBI (NCBI), yet all must pass through the CBIC, the effort truly results in little value-added. OPPT *must* be a good steward of industry data, particularly when the data are confidential, but technology can enable this without such a significant investment in resources.

RECOMMENDATIONS

The primary requirement that emerged from interviews and "brainstorming" conducted with EPA personnel was the need for greater access to TSCA information. This includes not only EPA access but access by industry and the public as well. In order to optimize the management of TSCA information, OPPT must make improvements in the receipt, processing (actual analysis and workflow) and distribution of that information. First, EPA must pursue electronic receipt of TSCA submissions by working with industry to establish standards (or standard formats) for submission, and then encourage their use with incentives, if necessary. Second, EPA should establish a single core TSCA data store that can be accessed and used simultaneously for all TSCA functions depending upon unique requirements. Finally, EPA must make the information available, real-time over a widely used medium such as the Internet. A well-implemented architecture based on these principles would result in much faster, cheaper, and more secure processing. We make specific recommendations that are addressed in the following sections under the headings Receipt, Processing, and Distribution.

Assign Program Management

Before any of these changes can take place, however, program responsibility must be assigned for sections 4, 5, and 8 processing. This is a fundamental recommendation. Some organization must provide a coordinating role before these changes can be implemented. This does not necessarily mean reorganizing OPPT. An existing branch can be given the responsibility for overall coordination including process improvement projects. If these responsibilities already reside with a specific branch or individual, management should make this clear to OPPT as a whole. Whoever is assigned this function for each submission type becomes the data owner for those submissions. The data owner then has the authority and responsibility for the quality, timeliness, and availability of that

data. Support services, including information and records management functions, must be responsive to the data owner.

Receipt

While data receipt is the point where EPA comes into possession of the submission and before it begins analysis, we use the term to include submitter actions as well. The optimum data exchange process begins with the creation of an electronic record as far "upstream" in the process as possible. For a regulatory reporting process such as TSCA, this begins with the submitter. In the current process, however, submitting TSCA information today is likely one of the few reasons a company still needs a typewriter. For the simple reason of storing this information electronically, as with most of its other business documents, the company has reason to "go electronic." Electronic receipt of the submission provides both direct and indirect benefits to EPA. Directly, it eliminates the need for several data entry contracts OPPT is currently funding (as well as internal data entry) and the attendant data entry errors. Indirectly, it allows for more process reengineering which leads to even greater savings and more effective processing.

Establish Electronic Reporting standards

EPA must establish a standard format for TSCA submissions, ideally with industry input and agreement. A standard format should be consistent, to the extent practical, with other efforts in EPA and the government as a whole. EPA should provide guidelines for electronic reporting methods. For this reason, it is useful to bisect data submission processes into two types. The first of these types may be characterized by processes which are transaction-oriented (as opposed to textual or narrative), simple in data content, routine, and high volume. The result of the submission may be automated processing by database or data propagation to other databases. In this case electronic data interchange (EDI) may be used by OPPT in the future once the central TSCA database is implemented. A good example of this type of submission is the Manifest Reporting process under the Resource Conservation and Recovery Act (RCRA) of 1976. It could also be used to link international databases.

The second type of submission process may be characterized by submissions where the data set is less-strictly defined, the format is more textual or narrative, the data content complex, and the volume of submissions relatively small. This type of process is also characterized by the submission of several forms and supporting documents. This process, in contrast to the first type, requires analytical processing by a subject matter expert (SME). Processes of this type, which include TSCA, are usually best served by an Internet-based, document management approach, rather than EDI.

TSCA is not unique from other EPA data collection processes in that a large part

of the submission is in the form of supplemental material such as laboratory studies. In many ways, the biggest challenge for EPA is to establish standards for the electronic submission of these forms and documents. One option is to make extensive use of the de facto standard known as Portable Document Format (PDF) popularized by Adobe, Inc. PDF files are images that have a variety of value-added features including the use of form fields, word searching capabilities, and digital signatures. Adobe has provided the specifications for PDF as an open standard, and a limited number of companies have made use of this specification. In this regard, while PDF is in fact dominated by Adobe, it is not wholly proprietary.

Government agencies are increasingly using PDF as a way to store business documents. For example, the U.S. House of Representatives Law Library makes the U.S. Code and other documents available over the Internet in this format. OPPT should define the formats which are acceptable for document submission and storage. This decision should be based on the need to:

- u manage documents as documents for ease of reading rather than as database records;
- u make sure the data storage format is not destined for quick obsolescence;
- u keep the data in a format that does not require exotic or unique software to view; and
- u be able to convert the data to another format should that be required.

Using these criteria, PDF excels over most other formats. Because it is easily converted to and from other image formats, relies on an open specification, and is so widely used, PDF is a safe choice that is not destined for obsolescence in the near future.

New functionality available in Adobe Acrobat 4.0 also makes PDF a good candidate format for receiving and storing TSCA submissions. Released in March of 1999, Acrobat 4.0 provides an easy method for attaching the digital signature to the actual submission, be it a PMN or a laboratory study. Signatures can be validated within a public-key infrastructure (PKI) in real-time or locally with a public-key "address book." Either way, this capability ensures authentication and non-repudiation. It also ensures data integrity, as any changes after signature are detected through a hash algorithm. This is addressed in more detail in Chapter 5.

PRESENT "SINGLE FACE TO INDUSTRY"

EPA should establish, at the Agency level if possible, a single submitter

registration authority, such as the Office of Environmental Information (OEI), to avoid redundant effort and industry confusion. This authority would establish and manage the broad procedure for electronic reporting that would include assigning external user identification numbers and passwords, establishing acceptable security and encryption standards, maintaining master lists of submitter information, and coordinating with other government organizations for consistency. Establishing these procedures is crucial to electronic reporting and will ensure a smoother transition for the entire Agency.

The current timeframe for implementation is set for the year 2001 or beyond. Given this lag, OPPT must establish its own central registration process for all TSCA data if they wish to move forward. This should at least be coordinated with OEI such that the data collected from submitters are generally consistent with OEI's plans for central registration. In this way the data OPPT collects from submitters could conceivably be migrated into a larger registration system at the Agency level.

MAKE MAXIMUM USE OF INTERNET CAPABILITY

The Internet is a tremendous resource that can impact all phases of the process. Compared to other electronic media, it is the fastest, cheapest, most intuitive, and most secure approach.

While issues of data security and CBI are real, the use of web-based submission architecture is potentially much more secure than the current media. The current standard for secure web sessions, Secure Sockets Layer (SSL), uses robust encryption algorithms (128-bit or more) to encode the back and forth data exchange between host and client.

MAKE THE BUSINESS CASE FOR INDUSTRY.

Even though there is growing interest, and impatience, within industry to convert to an electronic process, a business case must still be made for implementation. While the case may be exceptionally straightforward when the benefits are obvious and the costs to implement low, such as data entry submission over the World Wide Web (WWW), requirements to purchase software or program applications necessitate a greater "marketing" effort. In these cases the benefits to industry of the electronic process must be explicitly identified. Of the three submission processes, the PMN is where manufacturers may see the greatest opportunity for cost avoidance and thus be most interested in using. It may also, however, require the greatest effort on their part to implement. To encourage participation, the EPA should consider offering financial incentives (e.g. waiver of submission fees, either completely or only slightly), compress product approval time frames where appropriate, posting electronic status of PMN's for company review on a web page, and, or increasing fees for continuing paper submissions.

While a few companies may willingly convert to an electronic process even in the absence of a business case, EPA can move away from paper submissions much faster and more completely than without giving companies the incentive.

Reengineer CBIC Processing

EPA may not, in the foreseeable future, receive all TSCA submissions electronically. To the extent that it does, however, the CBIC role can be greatly reduced. The purpose of the CBIC and CBITS is to control and track submissions within EPA for security purposes. For electronic submissions, little to no data entry will be required. In addition, the tracking functionality currently satisfied by CBITS should be designed into the applications used to receive and store the submission. Electronic control of CBI records is potentially much more secure than the current process of distributing serial-numbered, hard copy documents that must later be collected and destroyed. Finally, routing through the CBIC delays actual processing of the submission for several days. OPPT should use the secure applications themselves to track the submissions. The CBIC should be used to process and digitize paper submissions. The CBIC would also perform the initial review for items received electronically before electronic routing to the SME.

Processing

Processing is the value that EPA adds to the submission after receipt and prior to distribution to the public. In the case of the Export Notice, this is cursory and mostly automated. By contrast, the Health and Safety, HPV, and PMN submissions can result in a detailed, highly analytical review. The following recommendations focus on better internal data management to improve processing.

Establish A Connected Core TSCA Data Store

Many of the data elements from the three primary core TSCA processes are the same, and the resulting output from each process similar. In order to reduce data entry, data maintenance costs, and improve user and management access to TSCA information, we recommend establishing a single connected core TSCA data store. Currently, Health and Safety data are stored in up to five locations, Export Notices two more locations, and PM N data in dozens of locations, none of which holds the entire record. NCIC information is stored in over one hundred individual files. It is very much impossible for OPPT or OPPTS managers to have real time access about the state of TSCA processing. Overlapping or redundant assessments are performed on the same chemical by different branches. A single connected robust data store, using the Agency standard, Oracle, could easily manage the input, processing, and retrieval for all three processes. This enables universal access to accurate data and eliminates the primary reason for the

existing proliferation of independent applications and databases. A single Oracle data store will also lend itself to data sharing with other agency's and country's data stores. In addition, a single data store based on each chemical would provide a "life cycle history" of that chemical's body of knowledge. This concept is discussed more thoroughly in Chapter 3.

Develop Workflow Applications as Needed

With an Oracle back end database, the Information Management Division (IMD) can generate front-end applications for workflow and processing from such programs as Lotus Notes and document management systems. In this sense, the database record is *always* the official record, and applications simply become different ways of *viewing* the data. Lotus Notes excels at rapid development of these kinds of applications, and the applications can quickly be changed to meet existing requirements. The data, however, stay the same and never need conversion. In the current environment, the data are wedded to the application, and updates require the painful process of data conversion and normalization.

Workflow management can significantly improve internal processing.

- First, it automates the document routing process based on established business rules. By automatically notifying persons responsible that action is required on a document, the friction normally associated with inter or intra-office processing is reduced.

- Second, it enhances access to the information. A well-designed workflow application allows users to work off the same document enabling parallel, rather than serial, processing to take place.

- Third, it normalizes the processing such that relevant performance measures can be established and measured.

- Fourth, it gives decision-makers greater visibility of the process. Managers are automatically notified when action is required or processing has fallen behind schedule. While not always appreciated, it has the effect of holding individuals accountable for deadlines.

Distribution

OPPT must make greater use of the Internet as a means for the public to locate and retrieve TSCA information on a real time basis. Implementing the recommendations for receipt and processing, particularly creating a TSCA data store, would enable easy, direct publishing of "CBI-sanitized" documents to a medium such as the Internet. The Internet's democratizing effect helps ensure that the information is available to more than just those experienced in accessing

it, such as interest groups. In addition, this would greatly reduce the reliance on the NCIC, thereby saving more resources.

BENEFITS

OPPT has a tremendous opportunity to realize significant efficiencies from these processes at the same time that it is making the information more valuable to EPA employees and to the public. Reengineering benefits traditionally fall into two categories: direct benefits, which are actual dollar cost savings resulting from, for example, the elimination of data entry contracts and indirect benefits. Indirect benefits typically include all the efficiencies associated with better, more timely information. While difficult to quantify, studies have estimated the value of indirect benefits at three times the direct benefits. We have identified both direct and indirect benefits available to OPPT should these recommendations be implemented.

Cost Savings

The primary direct benefit is cost savings through the elimination of the data entry function. Other functions which can be eliminated in the electronic environment include document control and data conversion functions. Currently, on three contracts, OPPT spends approximately $1.8 million a year on these activities. The bulk of this is spent on the operation of the CBIC. The role of the CBIC is to receive information that may be confidential, route it to subject matter experts, enter certain data elements for tracking purposes, make copies, and maintain the archive of CBI files. In theory, none of this would be necessary where electronic submissions are received and routed automatically to the proper destination. In practice, paper may not be eliminated in the near future, but it is reasonable to believe that 50% of submissions could be received electronically in the next two years for all submissions. It is also reasonable to believe that a commensurate savings of nearly one million dollars a year could be negotiated based on this reduced workload.

In addition to the contract savings, EPA could realize savings simply from the consolidation of various databases. The consolidation of TSCATS, CECATS and Triage databases into TSCAT 2.0 was an essential first step in the creation of a single core TSCA database. Consolidating the three systems has saved resources, maintenance, and population. It is much more efficient to consolidate the information to the maximum extent possible as opposed to having fragmented information. Unfortunately the estimate of potential savings is impossible to derive but is likely significant, given the dozens of applications that have proliferated throughout OPPT over time.

Indirect Benefits

Faster Processing

One of the primary benefits that OPPT can realize from a central database is much faster processing of records than is currently possible. This would impact tremendously on the body of knowledge used to study these substances. In addition, it is impossible to determine the number of studies that are performed which may have otherwise not been warranted due to the existence of an earlier study. Also, automated receipt and processing would ensure prompt notification, posting and processing.

Better Data Quality

An often overlooked but still important benefit is the fact that electronic receipt, processing, and distribution of information can virtually eliminate data accuracy problems. In addition, data aggregation into a single data store eliminates the version conflict and configuration management problems that occur when storing the same data in more than one location. GroupWare applications such as Lotus Notes excel at resolving these conflicts.

Finally, data normalization is often a problem when the data are stored in various locations. Examples include date field formats (mm/dd/yy versus ccyy/mm/dd) and text versus numeric field storage of the same number. Building a single data store requires the creation of an aggregated data dictionary. The effort required to develop the data dictionary is rewarded through better data quality and a more thorough understanding of what data are being collected and maintained.

Better access to data for management use

Process control is the act of gathering information about a process in order to determine whether or not the process is operating effectively and efficiently. Data are collected and compared against certain parameters and adjustments made as necessary. Currently OPPT management is severely limited in the amount and currency of data available for evaluating TSCA in all aspects, including both internal processing as well as its effects within industry including costs and compliance. The lack of a single data store of information makes this kind of process control nearly impossible. Aggregating the TSCA information into a single data store would provide management with the ability to perform ad-hoc, "snapshot" analyses to determine, for example, current processing workload and throughput or the existence of trends among TSCA submitters.

Obstacles to Implementation

CBI AND THE CBIC

Based on our observations of OPPT practices, individuals are laudably dedicated to the concept of being a good steward of industry CBI. This presents a challenge to process improvement, however, when CBI is used as reason against change. Our observations also indicate that the actions of the CBIC and population of the CBITS in some ways almost take precedence over the primary goals of TSCA. For example, we have heard the statement that CBITS is the most important application within OPPT. While keeping track of all data is crucial, tracking submissions is a support function and can be incorporated into any application. The fact is that OPPT can maintain at least as much information security as is currently possible in an electronic environment that includes submission over the Internet. Overcoming perceptions to the contrary requires most of all a commitment to change from management.

CULTURE

OPPT is faced with the challenge of overhauling its entire TSCA processing architecture. The biggest challenge to replacing the "old growth" of existing systems and processes is, as in most IT projects, making the commitment to change. OPPT personnel recognize the need for improvement and welcome its arrival, but also need to be convinced that a specific change is the "right direction." The commitment must come from management and must be demonstrated by follow-through. Based on interviews with OPPT personnel, the follow-through is what has been lacking on previous projects.

INDUSTRY

Because electronic receipt of the data enables many more efficiencies, industry participation is critical. In addition, electronic submissions are widely recognized within industry as long overdue. There are two requirements to ensuring industry participation in this effort. The first of these is to illustrate the business case as described previously in this chapter. The second is to alleviate concerns that may exist about electronic management of CBI. The technology exists to preserve the confidentiality of all information in a manner much more secure than in the current process. EPA must be prepared to convince industry that it has a comprehensive strategy for managing this information securely.

ORGANIZATION OF THE REPORT

This report documents a systematic approach for the EPA to implement the capability to receive, process and distribute TSCA data electronically. Chapter 2 discusses the procedures under which electronic submissions can be implemented with minimum investment on the part of the manufacturer. Chapter 3 describes the electronic architecture and workflow for managing and storing TSCA data

electronically. Chapter 4 addresses public access of TSCA data over the web. Chapter 5 identifies migration steps and issues that need to be resolved for the EPA to transition to electronic management of TSCA data. Finally, chapter 6 discusses the management level roles and responsibilities to effectively manage TSCA data in an electronic environment.

Chapter 2
Establish Electronic Reporting

OVERVIEW

Electronic reporting enables greater process improvement in all areas of TSCA processing. The three processes covered by this study, Export Notices, Health and Safety studies, and PMNs are all prime candidates for conversion to electronic submission. This chapter describes procedures under which electronic submissions for each process can be implemented quickly, effectively, and with minimum investment required on the part of the submitter. Public access considerations are discussed in Chapter 4.

GENERAL PROCEDURES

The Internet provides the backbone for electronic submission of TSCA data. Because the Internet is standards-based, excels in document receipt and publishing, and is ubiquitous and intuitive, the World Wide Web provides the ideal medium for TSCA data submission and distribution.

With the exception of the Export Notice, TSCA submissions can be characterized as non-routine, complex, and text and graphic-intensive. Because of this, Health and Safety submissions and PMNs are best managed as documents, rather than transactions. The most effective document submission process includes a submitter using a web browser to fill in basic, *meta* data about the submission with the user then attaching a file constituting the actual submission. Using the Health and Safety process as an example, a submitter would browse to the TSCA submission web page and fill out a hyper-text markup language (HTML) form that mirrors the Health and Safety Data (HASD) cover sheet. The HASD cover sheet is a voluntary industry submission that identifies meta data which is used to index the Health and Safety study in the TSCA Test Submissions (TSCATS) database. Upon completing the cover sheet, the submitter would then include the actual study as a file attachment. While variations of this procedure introduce some complexity to the process, the basic concept is very simple. PMNs can be submitted in essentially the same manner, with a different meta data form. Export Notices, which require no study or form, will require the submitter to fill in only five or six simple data elements in an HTML form.

Central Registration and Receiving

In order to reduce the overhead associated with managing electronic submissions,

for both EPA and the submitters, OPPT should centrally manage access and submission of TSCA data. The first of these issues, access, requires central registration. OPPT must create and manage an authorization control list (ACL) of submitters of TSCA data. Through the mechanism of central registration, a submitter will require only one user name and password in order to submit and access their data. With a broader scope, this should be coordinated with efforts that are taking place Agency-wide, headed by the Office of Environmental Information (OEI). At some point these efforts may merge or at least coordinate. Figure 2-1 illustrates a web page that could be used for centrally registering submitters and linking to the appropriate HTML form for submitting TSCA data electronically. Central registration is discussed in more detail in chapter 5.

Figure 2-1. TSCA Central Registration and Receiving Page

The second issue, central receiving, provides a single point of entry for web submissions. The submitter must only access a single web page in order to be able to submit TSCA data, whether those data fall under sections 4, 5, 8, 12 or HPV. The combination of the two concepts creates the "single face to industry," at least at the OPPT level, an important element in industry's acceptance of the

electronic reporting.

In addition to providing the "single face to industry," the central receiving point can reduce internal processing overhead. For example, functions such as error checking and maintaining transaction logs for web submissions can be centralized.

Processing Upon Receipt

The CBIC is assumed to be the point of receipt for TSCA data submitted by industry. Export Notices, because of their simplicity, should present few if any receipt processing issues. Health and Safety and PMN submissions, however, can present some issues, and we describe a practical solution for managing these submissions in the following paragraphs.

Figure 2-1 illustrates the receipt processing architecture. Individuals at the CBIC will use a graphical data entry and review screen tied to the TSCA data store. If a paper submission is received, the CBIC must convert the paper to an electronic PDF format by processing through a high-speed scanner. Current high-speed scanning technology can convert a piece of paper to an image in less than three seconds. As an example, a typical submission that may total 100 pages of information can be scanned in less than five minutes.

The resulting file must then be uploaded into the TSCA data store and indexed with basic information such as date received, submitter, and presumed type of submission (PMN or Health and Safety Study). In the case of the Health and Safety Study, the individual processing should check that the study is not associated with a HASD cover sheet entry previously submitted electronically. If so, the study file should be appended to the HASD cover sheet record. The paper document should then be archived.

PMNs may also be submitted on CD-ROM, particularly if the submitter does not place trust in the confidentiality of a web submission. These PMNs must be uploaded at the CBIC. First the CBIC would scan the CD-ROM for viruses and then save the file to local archive. Then the individual would create a record in the application with the basic data set (date received, submitter, presumed type of submission, etc.). The individual would then attach the file to the record and upload the submission into the TSCA data store.

Web submissions require a cursory review. An individual from the CBIC should browse through new submissions and ensure that the submitter has completed the very basic data. Some of the error checking can be performed by the web application, but the CBIC should still ensure that something intelligible was submitted.

After the CBIC has uploaded the submission into the TSCA data store, the individual would electronically "route" the submission into the Document Control Office (DCO) based on the submission type (PMN or Health and Safety). The DCO would then review, add some information to the document record (such as CBI content, actual submission type, chemical, etc.) and then "route" to the next phase of the workflow, which, for example, in the case of PMNs, would be the Initial Review.

It should be noted that a security plan for the submission, receipt, and processing of CBI should be developed by the EPA to ensure proper treatment of sensitive data in an electronic environment.

Figure 2-1. Receipt Processing Overview

SPECIFIC SUBMISSION PROCEDURES

Export Notice

TSCA Export Notice processing offers significant opportunity for process improvement. While OPPT does not expend tremendous resources to process the 10,000 notices received annually, baseline review of the process suggests more effort than necessary is required to process 12(b) submissions. In addition, current processing may inhibit OPPT's ability to meet the 5-day requirement for notifying the country of import.

The Export Notice is well suited for exporter submission over the Internet. Use of an HTML form offers the ideal solution for both the manufacturer, who would no longer have to go through the process of sending a formal letter, and EPA contractors who no longer would be required to sort and enter the data into an application. Specific characteristics of the process supporting this conclusion include:

- ◆ The data set is small and well defined. A simple HTML screen can allow the exporter to type the information on the screen and submit it directly to the EPA server.
- ◆ There are no graphics or attachments included with a normal submission.
- ◆ The quality of data entered can be controlled through drop down boxes, JavaScript and other HTML tools.
- ◆ The regulation does not require a signature on the submission.
- ◆ The majority of submissions are not business-sensitive (it is NCBI).
- ◆ Direct entry into the EPA server can dramatically speed up a process, which has a narrow 5-day processing window.

While the data set can be formatted by other means, the other solutions are not ideal. For example, e-mail is not an ideal choice because many submissions may not conform to the proper file specification, and attempts to automate the upload process may populate the database with irregular data. Manual intervention and data normalization would be required. An electronic form or template would not provide much value because of the small size of the data set, and because there is no existing 12(b) form that manufacturers are familiar with.

ELECTRONIC SUBMISSION SCENARIO

A manufacturer intending to export a chemical substance that requires notification must inform EPA of its planned shipment at least seven days before actual embarkation. It must identify to EPA the following:

- Name of subject chemical,
- Name and address of exporter,
- The country of import,
- The date of export or intended export,
- And the applicable section of TSCA.

Instead of drafting a letter, someone at the exporter's office would access the EPA electronic submission web site. After logging on with the organization's user name and password, the web site would direct them to the TSCA page and, from there, to the 12(b) submission page illustrated in Figure 2-3. At this point they would be given the option of reviewing previous entries they had made or submitting a new notice.

Using HTML entry "boxes," the user would fill in the applicable data elements necessary for submission. For data elements such as country of import, it is desirable to control the user's input, and thus preserve data quality, through the use of a drop-down box or other tool. The same logic can be applied to any other data elements where there is a finite choice of entries, such as submission type, for example. The most cost-effective application of this concept is to use a tool such as JavaScript. JavaScript is a programming language that can be embedded in HTML forms to provide further value to the screen. It is very simple, for example, to program a JavaScript routine to validate the CAS number check-digit and warn the user when the CAS number appears to be incorrect.

After the user is satisfied with the data, the user would click on the "Submit" button. Via Hyper Text Transfer Protocol (HTTP) the data are sent to the web server and database. From there the application could generate both an e-mail confirmation to the manufacturer and an e-mail notification for the embassy of the subject country.

Figure 2-1. Export Notice Entry Screen

Once the data are submitted to the EPA server, they are resident in a TSCA 12b database tracking system. The existing system employed for tracking purposes, the Export Notice Tracking System (ENTS) is in Lotus Notes. Lotus Notes has the advantage of publishing directly to and from the Web.

CBI

Export Notices claimed as CBI account for approximately 25% of the total submission volume. Given computer security and encryption technology, EPA can establish the WWW connection using Secure Sockets Layer (SSL) and accept CBI submissions over the Internet in the same manner as the NCBI submissions. For legal purposes, EPA should provide a standard disclaimer saying, in effect, that the submitter understands and accepts the risk of submitting CBI data over the Internet, and, as an alternative, can instead continue to submit paper via registered mail.

PAPER SUBMISSION SCENARIO

It is reasonable to assume that the simplicity of electronic submission over the Internet, as described in the previous section, will result in a great reduction of paper submissions. OPPT must still be prepared to accept paper, however. The paper submission scenario would not differ greatly from the existing process. We believe, however, that the routing of paper submissions should be changed such that the data input group receives *all* paper submissions first, including those, which are CBI. Currently there is no benefit to routing the submissions through

the CBIC.

We recommend directing all submissions straight to the TSCA 12b center. This recommendation will necessitate a change to Agency rules that require the submissions go through the CBIC who keeps the originals and give a copy to the Hotline. The TSCA 12b center can make the determination, based on submission content, whether or not the item is CBI. The contractor will perform the data entry, including logging receipt, and can then stamp and forward the CBI submission to the CBIC for archive purposes only.

Health and Safety

Health and Safety studies are submitted under TSCA sections 4 and 8, For-Your-Information (FYI), and the Interagency Testing Committee (ITC). The Health and Safety submission is formatted by the individual manufacturer or laboratory submitting the information and is typically provided as a document or report. The length of the document can vary widely. Voluntary cover-sheets developed by industry and EPA help index the study by allowing the manufacturer to provide meta data which is used to index the study in the TSCATS 2.0 database. The study itself is archived, currently in microfiche, and, when needed, used in scientific analysis. Because the study represents the legal submission, EPA personnel do not modify it. EPA may receive 2000 submissions a year (and an average of 3.5 studies per submission). The business sensitivity of the information is at the discretion of the manufacturer but is limited by the statute. For example, about 50% of the approximately 400 submissions received annually under Section 8(e) are CBI. Typically only the submitter and chemical identity are classified as sensitive, and, on the cover sheet, the proposed use and quantity to be manufactured.

Health and Safety submissions are identified based on the reason for submission:

- Section 4. The EPA can require a manufacturer or processor to conduct testing of a chemical substance or mixture. Such testing is used to develop data otherwise unavailable with respect to the substance's potential health and environmental effects. The data are used by the EPA to determine if the chemical does or does not present an unreasonable risk of injury to health or the environment. Upon receipt of any test data which indicates that a chemical presents or will present an unreasonable risk, the EPA will, within 180 days of the date of the receipt of such data, initiate appropriate action under section 5, 6, or 7 to prevent or reduce such risk or publish a finding that such risk is not unreasonable.

- Section 8(c). Studies under this section are rarely received by OPPT. Section 8(c) requires companies to collect allegations of

significant, adverse reactions and keep records of them.

◆ Section 8(d). The EPA can require manufacturers and processors to submit unpublished health and safety studies on TSCA-covered chemicals.

◆ Section 8(e). Firms must report immediately to EPA any time they receive information that reasonably supports a conclusion that a substance presents a substantial risk to humans or the environment.

◆ FYI. Firms may voluntarily report study results to the EPA even if the study's conclusions do not warrant an 8(e) submission.

◆ Voluntary/ITC. The Interagency Testing Commission identifies chemicals that contain "significant gaps in knowledge" about their health and environmental effects. The ITC then requests that industry provide any unpublished data that may be available.

The Health and Safety submission is well suited for direct entry of cover sheet information over the Internet, with the actual test data submitted as an attached file. The HASD cover sheet is not prohibitively long and consists primarily of check boxes and short text box entries. Pilot projects have shown the average time it takes to fill out the electronic HaSD form is ten minutes. Test data are submitted as documents and are not intended to be altered or populate a database. Some 8(e) submissions do not include attachments so that the cover sheet constitutes the complete submission.

When both a cover sheet and attachments are included, the manufacturer's submission consists of two processes: filling out the cover sheet on-line and uploading the attachments. The manufacturer must fill out the cover sheet information required for indexing the study. If the manufacturer does not enter all the cover sheet data, the Records and Dockets Management Branch (RDMB) must complete it as discussed below.

The manufacturer would then attach the PDF file containing the digitally signed study or test data. (If the submission is an incident report or one-pager, it does not require a digital signature). After clicking the "Submit" button, the manufacturer would be taken to a "confirmation page" indicating that the transaction was accepted. The page would include the document control number (DCN) generated by the application when the submission is processed. Because the paper submissions would also receive a DCN, DCNs assigned electronically would be segregated numerically from paper submission DCNs. If the submission was not accepted, the page would indicate an error. The manufacturer would also be sent an e-mail confirmation with the DCN. OPPT would then have the indexing information and document stored in a database (currently TSCATS

2.0) for further review and dissemination.

USE OF PORTABLE DOCUMENT FORMAT AND ATTACHMENTS

Both the Health and Safety and PMN submissions may require file attachments. Because of this, EPA must select formats which are acceptable based on a number of criteria. We believe that the use of Portable Document Format (PDF) is a good choice for this purpose. Our recommendation is based on the following considerations:

- PDF is a de facto standard for storage of electronic documents both in and out of government.

- PDF is widely recognized in the Internet environment as an effective way to exchange documents, primarily due to the ubiquity of the free Acrobat reader.

- PDF preserves the integrity of the data since it captures an exact image of the document itself.

- PDF is platform independent. Acrobat readers are available for most platforms.

- Adobe Exchange allows the creation of electronic forms with value-added features. For example, JavaScript capability allows client-side data validation including mandatory field entry, arithmetic checks, and inter-field dependencies.

- Adobe Acrobat 4.0 supports multiple digital signature capabilities.

- Users with Adobe Exchange can export data into the Forms Data Format (FDF) which can easily be parsed and uploaded into a database.

- Development of each form is very inexpensive and can be applied across a variety of EPA applications.

Considerations for not choosing other options include the following:

- Application development can be expensive, slow, and process-specific. File sizes can be very large. Development also requires that EPA support the software after fielding.

- Word processing forms are software specific and not platform independent.

Because PDF is published as an open specification, it has become a de facto standard. While Adobe Inc. holds the major share for the PDF market, there are other products that support the PDF specification. In addition, many organizations are using PDF as a format for archiving documents with some legal basis. This helps ensure that the format will not become obsolete in the near future.

When designing a document management architecture, one issue that must be addressed is storage space requirements. With respect to the use of PDF as the preferred format for storage, we must note that the required storage space will vary dramatically based on the type of source document used to create the PDF file. Hard copy scanned into PDF can require 10 times the storage space that the same file would require had it been converted from another electronic source (such as a word processing document) using the Acrobat PDF Writer or Distiller products. Submitters should be encouraged to create PDF files directly from electronic documents rather than by scanning hard copy.

Signing the Submission

Because the health and safety study is the submission and requires a signature, the electronic process must be able to accommodate this requirement. Adobe Acrobat 4.0 can provide digital signature capability on PDF documents. We recommend that manufacturers be asked to submit the actual studies as PDF documents containing a digital signature. Because the laboratory study represents the actual submission, OPPT should create a PDF "signature page" with an electronic signature field. The submitter would insert the signature page as the first page of the study, electronically sign the field, and save the file. This would eliminate issues with regard to signing the HASD form. General issues regarding digital signatures are addressed in greater detail in Chapter 5.

Premanufacture Notice

The PMN is submitted to EPA on a form that has a fixed number of fields with well-defined data types. It is a relatively lengthy document that must currently be filled out in hard copy by the manufacturer. Completion by the manufacturer is likely to require analysis and input from several functional areas working collaboratively over time. The form may include attachments with additional information, such as a substance's physical characteristics, and may also include graphics such as molecular structure. Test data and/or health and safety data may be included with the submission as a separate document. EPA currently receives roughly 2000 PMNs annually, with the largest single generator submitting about 100 of those. It is normally considered a highly sensitive business document that requires special security precautions.

The PMN submission is the most complex of the three business processes, primarily because there are several issues which must be resolved before OPPT

can begin receiving and processing electronic PMNs. The PMN form can be entered over the web "on-line," but the form is of sufficient complexity that it is unlikely a manufacturer would prepare and submit the form over the web at the same time. Because of security perceptions with regard to secure sockets layer (SSL) encryption on the Internet, some manufacturers may also not feel comfortable submitting PMN data via HTTP. We recommend that the process satisfy "off-line" submissions while remaining flexible by preparing for the bulk of future submissions to be web-based.

The unique aspects of both the PMN data set and the process in which the PMN is filled out for submission require a specialized approach. This approach must be able to meet the needs of the submitter which include varying requirements for security and the need to work at-length on filling in the form. The latter precludes the use of an "on-line" entry and submission. OPPT's needs include automated data entry and the ability to manage the submission as a document and route it through an electronic workflow.

In order to meet the needs of the submitter, OPPT must create an electronic "tool" that allows the users to fill in the appropriate information, print or export the form, and submit to EPA either in paper or some format capable of upload into a database. The options appropriate for these requirements include the use of forms software (or software with forms capability) or the development of an executable application. Based on our analysis, we recommend the use of a forms-capable product. Specifically, we believe that the functionality available in Adobe Acrobat is a good match for this process.

The use of a PDF form gives the user an electronic template from which to work. Once the form has been completed, the user could have three options for submitting the form to EPA. First, the user can print the file and mail the submission to the EPA. This is appropriate when the user is not confident that an electronic transaction will be secure or when they only have access to the Adobe Acrobat Reader. The benefits to this process include the fact that it is easier for the user to complete, the completed form is more legible than hand-written forms, and the data validation improves the quality of the submissions. The disadvantage of this option is that the user cannot download the form and must therefore complete the submission in one "session". Another disadvantage is the inefficiency associated with the processing of paper versus electronic transmissions. These disadvantages are significant and argue for the user to obtain Adobe Exchange so that the submission can be downloaded and electronically transmitted.

The second and third options require the use of Adobe Exchange. In the second option, the submitter saves the file to a CD-ROM. This file, and any supporting files, is mailed to EPA. The third option is for the submitter to go to the TSCA web site, complete an HTML form containing a small number of fields necessary

to properly index the PMN submission, and then upload the PDF of the PMN (and any other supporting documentation) in a manner similar to that already described for Export Notices and Health and Safety submissions.

While there are three potential scenarios for electronic submission, they all begin with the submitter downloading the electronic PMN in PDF format from the TSCA web site. The manufacturer (or importer) then fills out the PMN, using either Adobe Exchange or the free Adobe Reader.

Scenario One: Use of Acrobat Reader and U.S. Mail

If the user only has access to the free Adobe Acrobat Reader, then they must complete the submission in one "session," because the file cannot be saved. When done, they would print out the hard copy, physically sign, and mail the form, along with other documentation, to the appropriate EPA address. Benefits of using this form include better data quality and more legible submissions, however, completing the form in one session is not practical.

Scenario Two: Use of Exchange and U.S. Mail

In this scenario, the use of Adobe Exchange allows the user to save the file while it is being worked on, and, most importantly, to save the file to a CD-ROM for mailing to the EPA. After completing the form, the user would click on the "Submit" button and save the file on a CD. This could also be used to automatically create a "sanitized" version by creating a new submission with CBI data deleted. In this case, the submitter includes two PMN files on the diskette (normal and sanitized version) and any supporting documentation in electronic format (chemical structure files, MSDSs, health and safety studies, etc.). These can be accepted in a variety of formats, including PDF, because they will be stored as objects in the database. The submitter then mails the CD-ROM to the appropriate address at EPA. Providing a digital signature, as an alternative to a physical signature, is discussed later in Chapter 5.

Scenario Three: Use of Exchange and WWW

This scenario is identical to the previous scenario with the exception the user would not mail a CD-ROM to EPA but would instead go to a PMN web page where they would upload the files directly, now to OPPT, but in the future to the data store after completing basic indexing data on an HTML form.

Signing the Submission

Because the PMN requires a signature, the electronic process must be able to accommodate this requirement. Adobe Acrobat 4.0 can provide digital signature capability on PDF documents. We recommend that manufacturers be asked to

submit the actual studies as PDF documents containing a digital signature. This is addressed more in Chapter 5.

Chapter 3
Establish Architecture and Workflow

To determine the best electronic architecture and workflow management system for receiving, processing, storing, and distributing data received from chemical manufacturers or resulting from internal EPA processes, LMI considered the key functional areas that OPPT processes support and the databases that accommodate those processes. In this chapter we identify and describe the key functional areas or categories of TSCA data, the primary databases that currently accommodate those functional areas, and three basic alternative electronic architectures for improving the management of that data. While we looked at actual data elements and their construction within databases, the purpose of our analysis was to identify high-level opportunities for strategically improving data management, not to recommend detailed database design structures and entity relationship diagrams. A detailed analysis of each database is necessary, but should be conducted as part of a comprehensive functional requirements analysis after data management strategies have been determined.

KEY FUNCTIONAL AREAS

As we began to investigate the key functional areas that OPPT processes support, it became immediately clear that OPPT has not formally identified them. Consequently, OPPT processes generate information output that is fragmented. We performed a high-level evaluation of the OPPT databases in order to abstract the key functional areas that the data appear to support. We identified four key functional areas. Those areas and a brief description are provided below.

Chemical Information

This functional area is concerned with discretely identifying and providing information about chemicals and active ingredients of commercial products. Examples of the data collected and stored under this functional area include Chemical Abstract Service (CAS) number, the 9th Collective Index (C.I.) Nomenclature, EPA Registry Name, and Synonyms.

Document Management

This functional area is concerned with identifying various characteristics which uniquely identify, describe, and monitor status of documents submitted to the EPA. Examples of the data collected and stored under this functional area include document type, submission stage, document identification number and CBI status.

Facility or Submitter Information

This functional area is concerned with discretely identifying the submitter of TSCA data and chemical production volume. Examples of the data collected and stored under this functional area include the submitter's formal name, phone number, abbreviated name, DUNS number, submitter identification, and production volume.

Test Information Type

This functional area is concerned with characterizing the various test information categories and providing test results data. Those categories are represented by the following examples:

Study Information Identification

Examples of the data in this category include title, study status, summary, study number and laboratory name.

Physical and Chemical Properties Identification

Examples of the data in this category include boiling point, density, flash point and explosive properties.

Environmental Fate

Examples of the data in this category include biodegradation data, biological and chemical oxygen demand, photolysis, and stability in soil and water.

Health Effects

Examples of the data in this category include acute inhalation toxicity, carcinogenicity, acute oral toxicity, and acute dermal toxicity.

OPPT Databases

We observed that processes within OPPT are not designed to receive, store, process and distribute information in a consistent and integrated manner. Resulting databases to store that information have by necessity proliferated and contain much of the same information but cannot communicate with each other.

Because the same data reside in various databases, albeit in different formats, platforms and applications, the data must be entered separately into each database.

We found that many of the databases support common functional areas, some are for non-confidential business information and some for confidential business information, separate and distinct databases exist for each section of TSCA, and some databases contain historical data that and are no longer updated because they have been superceded by other databases. In 1994, mainframe databases were converted from Adabase to Foxpro, however, some data were not converted and some that were are not currently accessible. Consequently, manufacturers must resubmit data that may already have been collected but is not retrievable. Therefore, valuable historical data is lost. Additionally, scientists often have obstructed access to important data. Notwithstanding the confusion that characterizes the current data storage situation, we identified several databases that we believe are critical to OPPT's mission and should be considered as top priority candidates for reengineering. Those databases along with the primary functional areas they support are identified in table 3-1.

Table 3-1

	Chemical Information	**Document Management**	**Facility/Sub mitter Information**	**Test Information Type**
TSCATS 2.0	X	X	X	X
CRTK	X	X	X	X
CORR	X			
CUS	X	X	X	
CRS	X			
CBITS	X	X	X	
NCIC Database	X	X	X	
CICIS	X		X	

CCID	X		X	
HPV Internal Status Report	X	X	X	
HPV Master Register List			X	
CORR	X			
12(b) Export Notification	X		X	
MEGA	X	X	X	X
DocLOG	X	X	X	X
MITS	X	X	X	X
Penta	X	X	X	X
CIMS	X	X		X
SIDS	X	X	X	X

DESIGN ALTERNATIVES

Several design alternatives were considered to facilitate receiving, managing, and distributing data within OPPT, including a data warehouse, linked databases, and a central data store.

DATA WAREHOUSE

A data warehouse is the collection of information in one system from various

enterprise systems, including external systems, that gives management a global view of information. With an enterprise view, management can make more informed decisions and ultimately take corrective or preemptive actions.

Creating a data warehouse, however, is a complex, daunting, and expensive task. Not only must the data warehouse overcome technical and political boundaries, the benefits of consolidated operations and maintenance are not realized. Furthermore, a successful data warehouse is largely based on the validation and integrity of source data. In fact, the Data Warehousing Institute estimates that addressing and resolving data integrity issues represents 70% of the cost of a single warehousing project. This estimate, coupled with serious data gaps, data redundancy and data integrity problems identified through the course of this study, suggests that a data warehouse is not the most appropriate solution for OPPT to consider.

Linked Databases

For those organizations that have sound data validation and data integrity practices, linking diverse, legacy systems can be a cost-effective alternative to centralizing data. If resources (e.g., time, capital or human) are constrained or management requirements do not dictate an enterprise-wide view to data, linking otherwise isolated databases can be an attractive alternative as well.

Like data warehousing, however, linking legacy systems does not eliminate duplicative operations and maintenance costs, including training, documentation, programming, trouble-shooting, help desk support, etc. Furthermore, integrating security and communications protocols across a vast array of legacy systems creates a complex and thus difficult (i.e., expensive) environment to support. Ultimately, because data validation and data integrity are still major concerns among a significant number of legacy OPPT systems, linking databases would only perpetuate these fundamental problems rather than correcting them.

Central Data Store

A central data store offers organizations the opportunity to consolidate resources and significantly reduce operations and maintenance costs over time. Typically, considerable upfront costs may exist due to new or enhanced requirement analysis, modified technical specifications, and investments in new hardware and software. However, the reduced long term costs and the benefits of centralized access, reporting, security, communication, administration, backup, etc. only strengthen the argument for a central data store.

Providing EPA researchers and scientists as well as the public with simplified and seemless access to chemical data is also of paramount importance to the EPA. To do so, appropriate resources must be allocated to effectively build and operate a central data store. In such an endeavor, the EPA would have the opportunity to

circumvent many of the data validation and data integrity problems of the past by integrating sound business practices for ensuring data quality within a new, central data store.

RECOMMENDED ALTERNATIVE - TSCA DATA STORE

Our primary recommendation for improving TSCA processing after receipt is to centralize core TSCA data into a central data store. We base this recommendation on several factors. First, from a good architecture planning perspective, it is much more efficient to populate and maintain a single database. The current state of fragmented storage is extremely inefficient. Centralizing data storage is not always the most attractive option, given modern database interconnectivity. In the case of OPPT, however, databases have proliferated to the point where some "pruning" and a lot of modernization are required.

Second, and more importantly, management does not have a good understanding of what processing is being done. In order for management to ask the "what if" and "how many" questions necessary to understand and improve processing, it must first have visibility of that processing. A data store which combines the three process similarities and data content can also provide management with a better understanding of where additional resources are required, and where they are available.

Third, from a research perspective, the lack of access to a central chemical "knowledge base" makes research more frustrating and difficult for the individual at the branch level. By incorporating data on individual chemicals from PMN receipt, health and safety studies, export notices, rule-making, and production data, as well as any other research information, the central data store can provide a "life cycle history" of each item on the TSCA inventory.

Perhaps most importantly, the central data store can enable easy distribution of NCBI chemical data to the public. This is consistent with, and essential to, OPPT's evolving role as information broker for TSCA information. We discuss methods for managing CBI data later in this chapter.

OVERALL ARCHITECTURE

In Chapter 2 we discuss the electronic submission of data. In this section we address the electronic architecture for receipt and management of TSCA data. Web submissions create a "network triangle" between OPPT, the submitter, and the EPA's web page hosts at Research Triangle Park (RTP). The same is true for publishing TSCA data to the web. OPPT must establish an architecture that connects OPPT with both the submitter and the public (for access), securely, through RTP. Figure 3-1 provides an overall look at the architecture. The geography of the specific architecture is irrelevant, with the exception of understanding how a submission may work. In this case, RTP maintains versions of the Lotus Notes/Domino application type for each submission (Export Notice, Health and Safety and PMN). In addition, RTP maintains an Authorization Control List (ACL) of registered users. When a submitter accesses the TSCA submission web page at RTP, they are challenged for a user identification and

password. Upon successful authentication, the submitter can then choose which type of submission to make. If the submitter chooses to provide an Export Notice, for example, they follow a link to that particular submission page, enter the data, and submit the data. The application generates an E-mail confirmation sent to the registered user's Internet E-mail address.

Periodically, perhaps daily, the RTP Lotus Notes application is replicated to the same application running on an OPPT Lotus Notes server located in Washington. The data are also passed through to the TSCA data store using some open-systems database compliant (ODBC) connectivity tool such as Lotus' Notes Pump. The data are now available for processing in OPPT. For public access, the process works in reverse. NCBI data are replicated periodically back to RTP where they are available for public access. In this manner, the public is not allowed access to the actual data record at OPPT.

Figure 3-1. Web Submission and Data Access Architecture

Notional Design

The most likely candidate system for a central TSCA data store is Oracle. Before any design takes place, however, OPPT must perform a thorough requirements analysis. From a high-level, we believe that the TSCA data store should initially focus on subsuming only the functionality of those databases as identified in table 3-1, which are core to the TSCA mission. Appendix A lists all the databases that

are currently maintained by OPPT and should be considered for integration with the data store at a later date.

Because this represents a major effort in reengineering TSCA information storage, the project must be undertaken with a long term planning horizon. A thorough requirements analysis should not only identify current requirements but also a reasonable anticipation of future requirements. We must emphasize, however, that this should not mean overloading the system design with "just in case" functionality to meet every user's perceived needs. Many systems development projects fail because of extra "nice to have" functionality that was never documented in the requirements. Often 80 percent of the development cost can be attributed to 20 percent (or less) of the requirements.

Figure 3-2 illustrates a generalized relationship between system complexity and the system's life-cycle cost. Integrated enterprise software that manages all the financial, accounting, materials and other similar information needs of an organization may be placed to the far right on the X-axis requiring a high degree of structure in a sophisticated relational design to manage all the data tables, elements, and relationships. These systems are relatively expensive to build, populate, and maintain. Document repositories, by comparison, are much simpler in design requiring a low degree of structure and as a result have a much lower life-cycle cost.

In a general sense, we believe that OPPT would fulfill its role as TSCA information broker with a system that features characteristics of both a document repository and structured database. We also believe that OPPT should begin the design from the document repository perspective and add structure as documented by the requirements analysis. An emphasis on a simple, yet flexible design may be the safest strategy.

Figure 3-1. System Design and Cost Relationship

Figure 3-3 illustrates an entity relationships model for the TSCA data store. This illustration depicts a chemical master table at the core of the database. Chemicals, in conjunction with a submitter master table, tie to submissions in one-to-many and many-to-many relationships. In addition, chemicals and submitters are related to information currently stored in the NCIC such as public comments on various rules. Miscellaneous information can be linked to each chemical to form a chemical library or knowledge base. The structure shown is also intended to be flexible. Other information can be added later based on that information's relationship to one or many chemicals. This may include, for example, risk management information. We must emphasize that this illustration is *not* a true entity relationship or database design document. It is intended to demonstrate a general approach to the data structure. The true design will be significantly more complex than shown, but, relative to most Oracle projects, most of the requirements could likely be met with a simple to moderately complex design.

Figure 3-2. Notional Entity Relationships in TSCA Data Store

Records Management

Records submitted by any person who manufactures, processes, or distributes a chemical or mixture that poses significant adverse reactions to the health of employees must be retained for a period of 30 years from the date such reactions were first reported to or known by the person maintaining such records. Records of adverse reactions to the environment must be maintained for a period of five years from the date the information contained in the record was first reported to or known by the person maintaining the record. In order to comply with this requirement, records in the TSCA data store must be assigned a record disposition category. The system should automatically date a document when it is saved as a record and preserve the date of receipt on records received. The date should remain constant without being changed when accessed, read, copied, or transferred. The system should look for records to be reviewed for disposition at time intervals specified by the administrator. When the retention review date is reached, the system should send a notification to appropriate reviewers and place the record in their in-box. Based on rules set up by the RDMB, records can either require or not require disposition approval by the reviewers. For example, if approval is not required and the reviewers do not take some action such as freezing a record, the disposition should take place automatically at the end of the review period. Users who authorize a record or set of records for destruction should be presented with a second dialog requesting confirmation of their decision. The system should provide a two step process for record destruction.

First, the system should identify and erase metadata records related to the files requiring deletion. Second, a clean-up utility should look for and erase content files that are not linked to a metadata record. The system should generate and send destruction notices to those responsible for records with specific information about the records and how they need to be disposed. The system should require confirmation that disposition has occurred as specified. Certificates of destruction could also be produced if required.

To provide for the most effective use of storage resources and enhance the overall system performance, the TSCA data store must provide automatic migration and calculation of migration candidate records based on complex criteria, including activity. The system should provide for the migration of electronic records to near-line and off-line media based on sets of rules defined by EPA. This helps to ensure that the most relevant and timely data will be physically stored so it can be rapidly accessed.

The TSCA data store should be backed up. The method used to backup the TSCA data store must provide copies of the data that can be stored off-line and at separate location(s) to safeguard against loss of records and other records management information due to system failure, operator error, disaster, or willful destruction. Following any system failure, the backup and recovery procedures provided by the system should provide the capability to complete updates to the data store. The system should provide the capability to rebuild forward from any backup copy, using the backup copy and all subsequent audit trails.

Other Considerations

While the TSCA data store is primarily for use within OPPT, consideration must be given for potential data sharing outside of the office. For example, OPPT's sister office, the Office of Pesticide Programs, has plans for developing a similar data store, in Oracle, based on the requirements of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). Common chemicals that inevitably overlap each law should link the "FIFRA data store" and TSCA data store. OPPT should perform a brief analysis to determine the degree of overlap and coordinate data definitions with OPP in order to provide a seamless link between the information stored in each office.

Similarly, OPPT should integrate its efforts with the international community to aggregate related chemical data. The European Community currently uses a system referred to as the International Uniform Chemical Information Database (IUCLID). IUCLID is used as a basic tool for priority setting and risk assessment for the European Union risk assessment program. The data is indexed by chemical and describes the effects of chemicals on human health and the environment. OPPT plans to share data with IUCLID and any U.S. versions of IUCLID that may be developed. Additionally, OPPT should plan for sharing data with other large external chemical data and hazardous materials databases such as EPA's FIFRA, and Department of Defense's (DoD's) Hazardous Material Information System (HMIS), Hazardous Substance Management System (HSMS), and the Environmental Reporting Logistics System (ERLS).

Secure Networking

To accommodate the need for public access as well as the need to protect CBI, and, in the future, submit CBI transactions, EPA should use two Notes/Domino servers. The first server would be configured to allow industry to submit information in a secure environment (using SSL v3 as the underlying authentication protocol). When accessing the first server, users would be challenged for an ID and Password that they would have had to obtain previously. SSL's public key technology (X.509) together with Domino security would then validate the credentials of the user and either allow or deny access accordingly. If this method of submission is not deemed adequately secure, the alternative is for OPPT to receive CBI data by paper or "off-line" electronic submissions such as on diskette or CD-ROM.

The second server would be configured for public access and would therefore not require an ID and Password to gain access. However, content on the second server would be limited to NCBI only, which would have been obtained via periodic one-way replication (or a push) from the first server to the second server. In this fashion, CBI information would never be open to public access.

Workflow and Processing

The TSCA data store aggregates and centralizes the data, but the real processing efficiencies cannot be realized without an intelligent way to access the information. Lotus Notes workflow applications, including several already developed, provide the interface and processing infrastructure. While the Oracle database represents the "back-end" where the data are actually stored, the Notes applications provide a "front-end" user interface. The Notes application also codifies the desired process.

OPPT's Chemical Information System (CIMS) uses a Lotus Notes Workflow management approach for activities involved in reviews and assessments for the existing chemicals program. CIMS provides an automated central repository for the creation, tracking, and sharing of existing chemical review documents. This database presents an integrated repository with links to testing actions, a chemical directory, and a document repository. A link to the TSCA data store is essential to provide central access to those data, further enhancing the productivity of EPA personnel.

Export Notice

Processing TSCA 12(b) submissions is relatively simple and can, to a great extent, be automated in a Notes application, including the automated generation of export notices to the countries of export. In addition, little data sharing or workflow is required relative to other TSCA submission processes.

Health and Safety

Basic processing of Health and Safety submissions includes performing a hazard screening assessment (health and ecological), setting priority for full risk screening, and performing a risk management analysis (RM1/2). The architecture and the roles and responsibilities in the reengineered process for each branch in

OPPT are depicted below in Figure 3-4. This illustration does not include public access to the data that is described in Chapter 4.

Figure 3-1. Health and Safety Processing Architecture

EXISTING CHEMICAL ASSESSMENT BRANCH (ECAB)

Once notified by the workflow application that a new submission has been received, the ECAB will view or download the submission in order to perform technical and scientific reviews. The ECAB will complete HTML screens designed for its role in recording and publishing the results of all submissions reviewed. Part of this effort includes providing further index information that may not have been submitted with the study. Once submitted, that data is a part of the health and safety data submission that can be viewed by the public according to the business rules to be established.

HIGH PRODUCTION VOLUME CHEMICAL BRANCH (HPVCB)

Once notified by the workflow application that a new submission has been received, the HPVCB will view or download the submission in order to evaluate submissions and perform hazard ranking, initial screening, and additional data development. The HPVCB will complete data entry forms designed for its role in recording and publishing the data it is responsible for developing. Once

submitted, that data is a part of the health and safety data submission that can be viewed by the public according to the business rules to be established.

As part of the OECD SIDS program, OPPT/Risk Assessment Division (RAD) is responsible for peer reviewing SIDS assessment reports generated by US manufacturers as well as reports from OECD member countries. This activity involves review by technical experts in RAD and Economics, Exposure, and Technology Division (EETD). The basic process involves receipt of the reports by RAD/HPVCB which then distributes the material to a defined group of technical experts in the two divisions. These experts provide comments through their management to designated staff persons in HPVCB so that the comments can be summarized and general conclusions can be reached concerning the quality of the reports. The review process usually requires the input from 8-10 technical experts spread among the two divisions.

CHEMICAL INFORMATION AND TESTING BRANCH (CITB)

The CITB monitors the workflow application in order to manage and report the status and progress on the processing of TSCA section 4 and 8d health and safety submissions. The CITB will complete HTML screens designed for its role in recording and publishing progress reports and final test data that are not already resident in the system. The workflow application can be programmed to produce various electronic or hard copy management reports for relevant parties on a periodic or as required basis.

Chemical Right-to-Know

The Chemical Right-to-Know (CRTK) initiative calls for the voluntary reporting of about 2,800 high production volume (HPV) chemicals over the next five years. The volume of reports to be submitted under this initiative represents close to about 50 percent of all current submissions to OPPT. While the processing of this information will be similar to the processing of health and safety data, the sheer volume of the data and the requirement to perform some unique preliminary tasks related to this data necessitate careful consideration as to how the initiative will be integrated into the reengineered business process discussed above. In this section we will address the preliminary tasks that need to be accomplished and then identify the roles and responsibilities of each branch in OPPT related to processing CRTK data.

Consistent with a fundamental principle of this study, we recommend that OPPT plan for capturing CRTK data with the intent of integrating them into the TSCA data store. Because the data collection process for CRTK is likely to begin long before any integrated architecture can be implemented, OPPT will have to develop or use existing automated information systems (AIS) for the CRTK requirements. The description in the following sections assumes the use of the integrated TSCA data store. Figure 3-5 illustrates the overall processing architecture, but for clarity does not include public access to the data.

Figure 3-1. CRTK Architecture (Macro View)

NOMINATION

The first step in CRTK is to get a commitment from a manufacturer of an HPV chemical for testing. The initial HPV list contains aprroximately 2800 chemicals. Under assumption that the manufacturers are willing to accept responsibility for their own "set" of chemicals, it is expected that the companies will agree to provide test data for certain chemicals by certain dates. Letters have been sent to those manufacturers requesting them to consider testing HPV chemicals they produce or import. This is the "Nomination" stage of CRTK. Data collection for this phase should include those companies or consortiums that have volunteered to provide data on specific chemicals or chemical classes, which letters have been sent, which companies or consortiums have responded, and what the management of each company or consortium has committed to provide in the way of test data. The architecture for electronic and paper nomination stage processing is illustrated in Figure 3-6.

Figure 3-2. CRTK Nomination Process

After receipt of the letter, the company or consortium representative accesses the CRTK web site and "signs up" to provide data for the subject chemicals. The representative provides data such as their corporate identity and the subject chemical or class. Hard copy responses by the representative will go to the RDMB. The RDMB will manually input data into the CRTK workflow application.

TESTING PLAN SUBMISSION/REVIEW

After a company or consortium has agreed to responsibility for some set of subject chemicals, it must submit a test plan to the EPA, to be accomplished through a third party, that describes what the company will do to fill in the voids in the Screening Information Data Set (SIDS). In the meantime, the companies or consortiums develop plans to determine what type of testing each chemical requires based on the end points defined by the SIDS. This will likely consist of only identifying the SIDS endpoints. This may include submitting an available but previously unpublished health and safety study and attaching it to the HaSD form. It may also include a plan to conduct a study where none has been performed or update an existing study based on new information. The plan itself may be submitted as a document. Data including planned completion dates are also submitted at this time. A macro-level depiction of this stage is illustrated in

Figure 3-7.

Figure 3-1. Testing Plan Submission

Similar to the nomination process, the representative either provides the test plan electronically or in paper form. Paper submissions go to the RDMB where accompanying data are entered into the CRTK application and test documentation is scanned and uploaded. Electronic submissions do not require processing by the RDMB. Notified of a received test plan, the HPVCB performs an evaluation. Deficiencies in the test plan as determined by the HPVCB may be resolved through conversations with the company or consortium, and the database is updated. If the test plan is initially or eventually acceptable, HPVCB marks the submission as "acceptable" in the CRTK application. When agreement cannot be reached, the HPVCB notifies the Existing Chemicals Branch (ECB) via the CRTK application that the chemical should enter the Test Rule process.

TESTING/PROGRAM MONITORING

During testing, the company or consortium may identify changes or problems with proceeding along the original test plan. HPVCB, with assistance from the various disciplines of the EETD works with the manufacturer to resolve any issues. The manufacturer should submit requests for any decisions or new information pertaining to any chemical to the web site for review by EPA.

DATA SUBMISSION

Submission of test studies is described and illustrated in Figure 3-4.

RM1 and RM2

The Risk Management 1 process is the first stage of the overall Risk Management function. It is designed to screen, prioritize, and select chemicals that appear to be of the greatest concern to human health and environment. This is generally limited to those non-polymeric chemicals on the TSCA Inventory that are produced in excess of 10,000 pounds annually. If additional testing is required, the chemical may then move to Risk Management 2 (RM2). The goal of the RM2 function is to identify options for reducing risk associated with a specific chemical. OPPT's workflow application for RM1 is part of the Chemical Information Management System (CIMS). It was designed to help manage the process flow described below.

RM1/RM2 processing involves several phases of analysis to determine the risk associated with the chemical. The RM1/RM2 processing includes input from the following functions: Chemistry Assessment, Production/Process Control Characterization, Hazard Assessment, Dose Response, Exposure Assessment, Risk Assessment, Economic Assessment, and Risk Management/Regulatory Control. Other inputs include data available from external sources and databases. As illustrated in Figure 3-8, each functional area adds value into the document record. The "box" contains the objects of the submission including an electronic copy of the submitted study, supporting documentation files, cover sheet information (for locating the record), and, depending upon the stage of processing, word processing files of the reports generated by the Chemistry Assessment function. While Figure 3-8 portrays this process in a serial manner, this is not necessarily how the actual workflow must proceed. Because all parties have access to the data at the same time, the serial nature of the workflow is only necessitated if input from one function is required before another function can commence. For example, because the Chemistry Assessment Report is a critical input to the Production/Process Control function, business rules require that it be processed serially before the Production/Process Control function. Therefore, after the Chemistry Assessment Report is posted to the document record, a message will be automatically sent to those in charge of Production/Process Control to begin that function. Figure 3-9 illustrates data sharing among all the branches that might require access to Health and Safety data in order to accomplish their respective functions.

Premanufacture Notice

Premanufacture Notices process in a manner very similar to the Health and Safety process described above. Once the PMN submission has passed through the DCO, it can be processed for content in the initial review. Managing the submission as a document allows everyone access to the record simultaneously for collaboration and coordination. Each subject matter expert (SME) in each functional area (such as Chemistry Assessment) can add specific information, or "value adding," to the document record. They can then "pass" the document record along to the next step in the process. In reality, the document is always available for viewing by anyone in the process. "Passing" the document is actually passing responsibility

to the next individual in the process by notifying them that it is their "turn." As illustrated in Figure 3-10, each functional area adds value into the document record. The "box" contains the objects of the submission including the PMN file, supporting documentation files, index information (for locating the record), and, depending upon the stage of processing, word processing files of the reports generated by the Chemistry Assessment function, hazard level assignment from the SAT, etc.

Figure 3-1. Electronic Processing in RM1/2

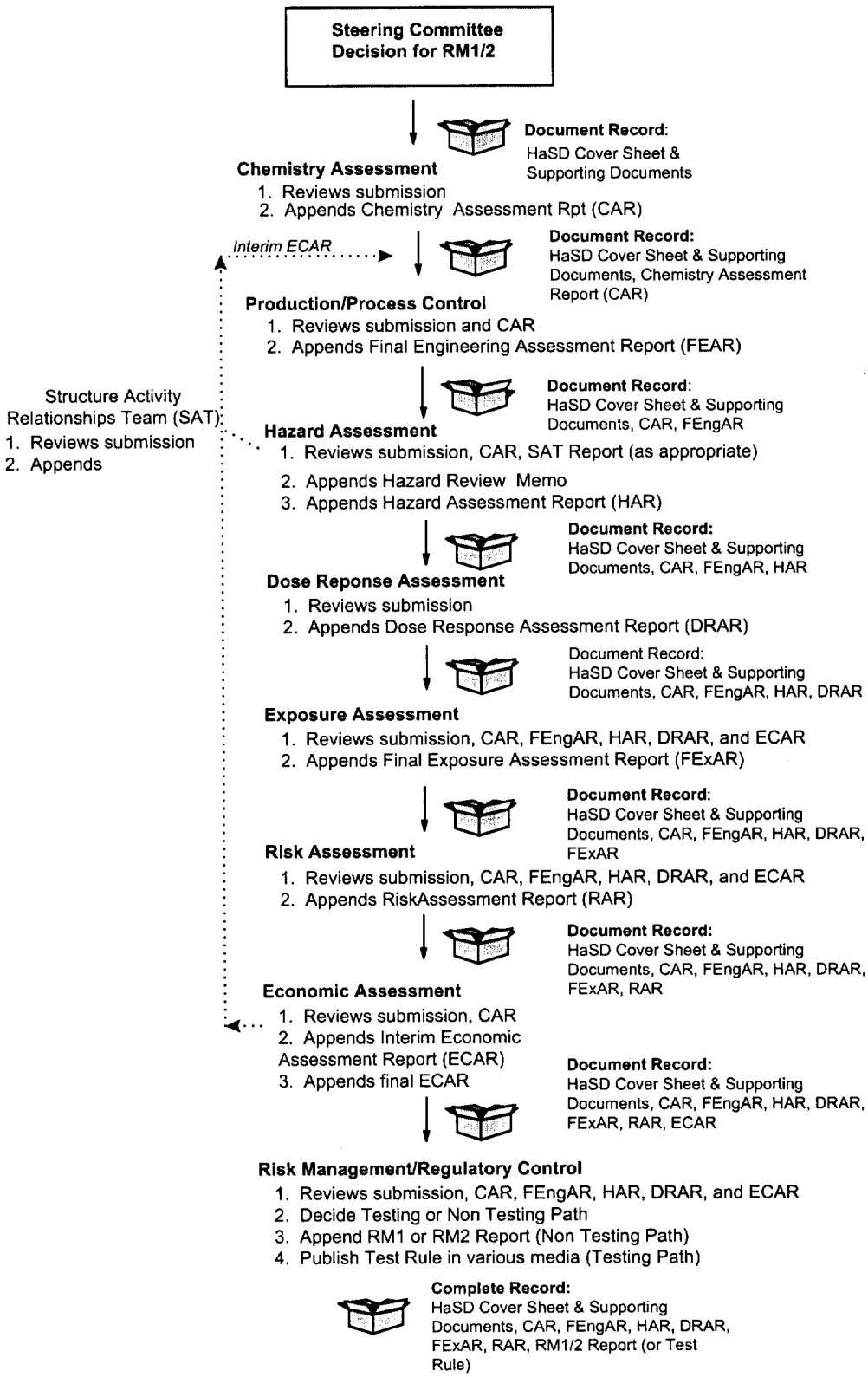


Figure 3-2. Data Sharing in RM1 Process

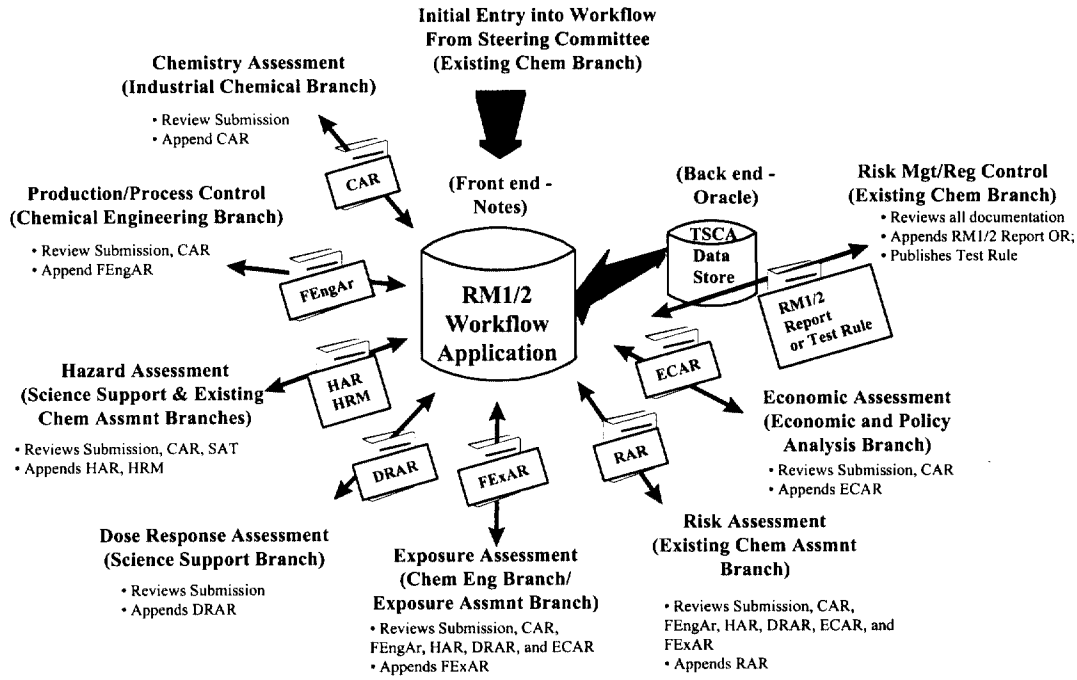


Figure 3-3. Electronic Processing of PMN

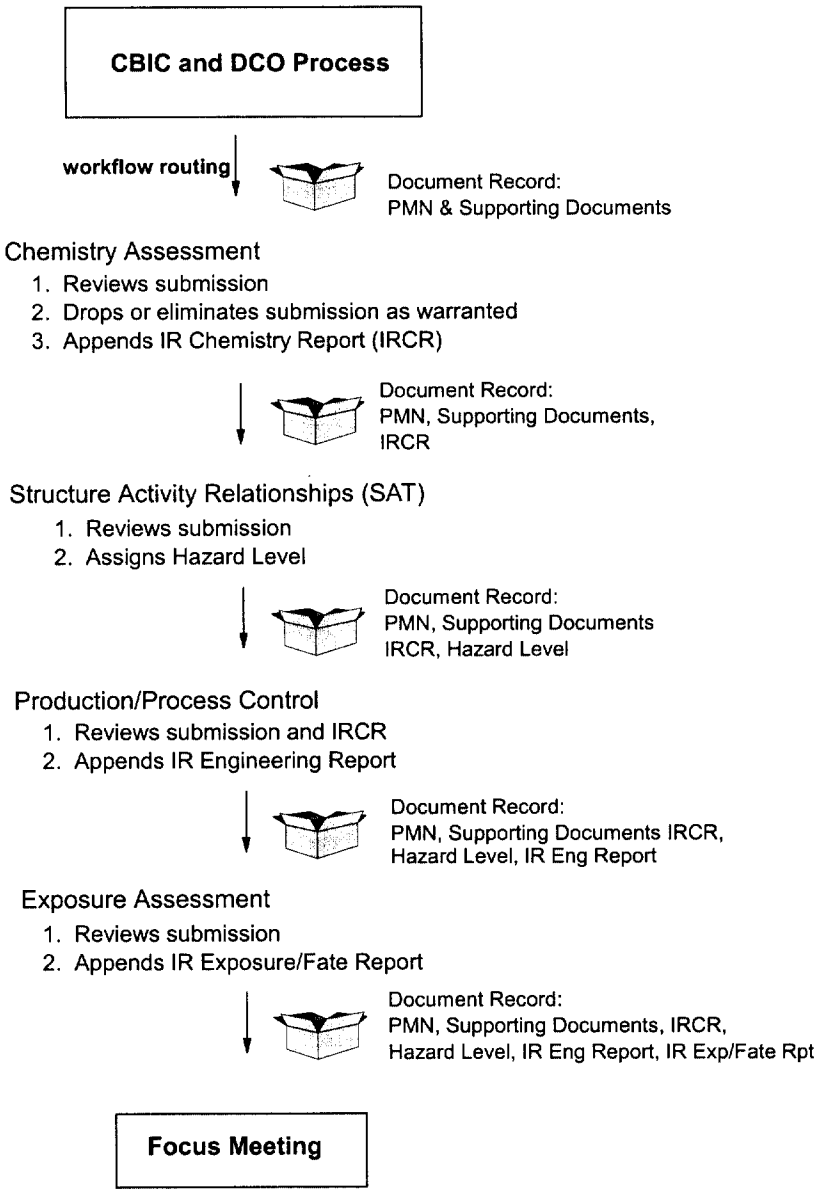
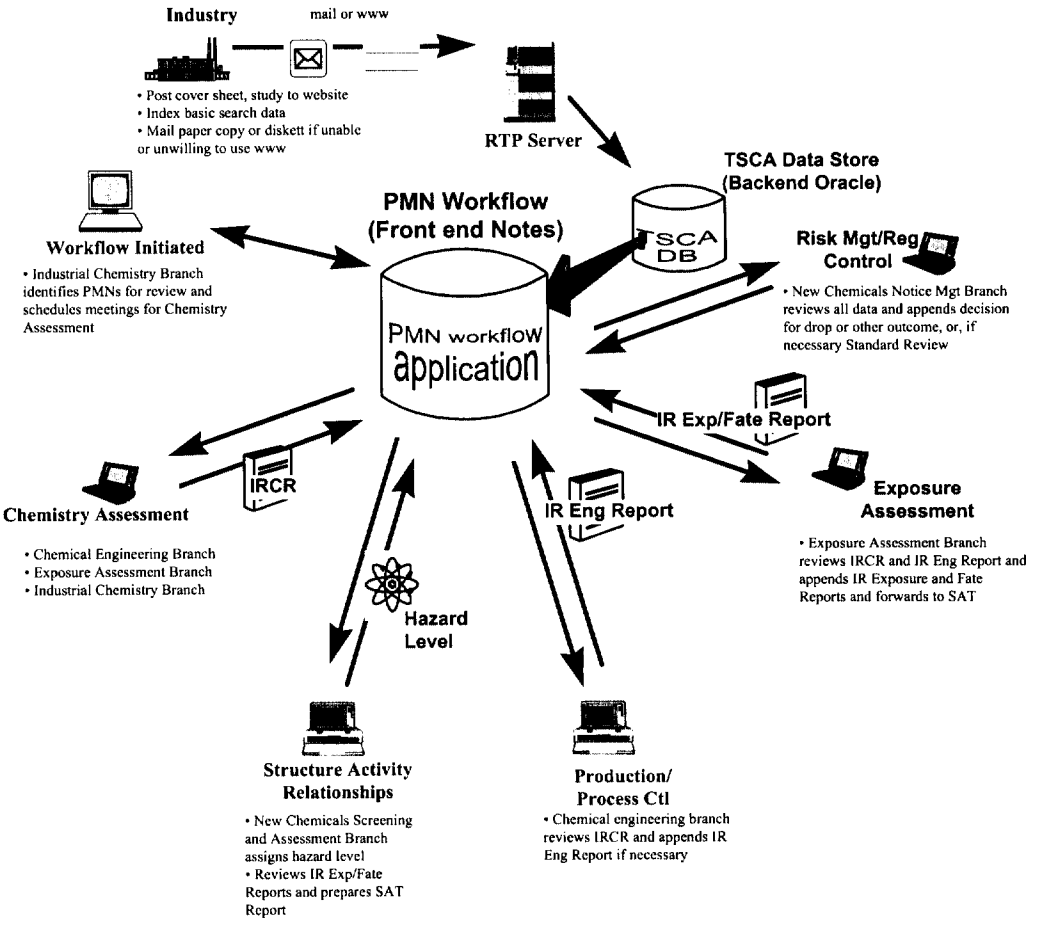


Figure 3-11 provides a functional view of this approach with emphasis on the common access to data. Again the process begins when the PMN is available electronically in the TSCA data store (the back end). The Lotus Notes workflow application is the front end used by all parties in the process to view and edit the submission. The DCO is then alerted to review the submission, adds value to the submission, and then alerts the next party in the process.

Figure 3-4. Data Sharing and Workflow During Initial Review



The Standard Review process would proceed along very similar lines, but would be more flexible based on the ad hoc nature of the Standard Review process. In any event, the workflow application should have the capability of defining what members of the team are involved in that particular review.

Chapter 4
Public Access to TSCA Data

In this chapter, we address issues associated with public access to TSCA data. The power of Domino, the Lotus web-publishing accompaniment to Notes, permits the direct publication of information to the web. Little to no HTML knowledge is required to construct a page which will intelligently, and dynamically, publish the contents of the database for web access. There are several general issues that must be addressed and resolved, however, with regard to web access. The foremost of these is "how much information is to be made available?" Obviously CBI submissions or data will not be accessible via the web. Generally, with regard to NCBI submissions, the Chemical Control Division (CCD) is responsible for determining and approving all information for public access as part of its responsibility for the risk management and regulatory control function. Issues associated with each type of submission are addressed in the following paragraphs.

EXPORT NOTICE

Distribution in this process consists of notifying the country of import as required by law. Countries of import can now be accommodated in three ways. First, they can continue to receive the letter that will be generated by the application. Second, they can choose to receive an automated e-mail notification. Third, they can also access a TSCA site listing Export Notice information. Because the Domino version of Lotus Notes enables direct web publishing, this requires very little effort to program. Letter notices can be eliminated if the documentation (such as a federal register notice) that frequently accompanies export notices is converted to an electronic format. For example, if a specific Federal Register Notice were stored electronically, the application could automatically determine if a notice was required to be sent to the country of import, and, if so, send an e-mail notice with electronic attachments (or references to a web page) to the country of import. In summary, the process could be completely automated.

HEALTH AND SAFETY

Electronic access to data is provided through query capability using specific search criteria. This access is illustrated in Figure 4-1. Management reports and data extracts may be produced electronically or by hard copy as desired. Tape or microfiche may also be produced as described above. Technically, the data from the workflow server would be replicated to a public access server at RTP. Only that data which are for public access would actually be replicated. CBI and other data that may only be for internal use would not be replicated, such that the data

set at RTP would be a subset of the actual database at headquarters. In this manner, the public would have read-only access to indexing information and laboratory studies approved for distribution, but the public would not have any access to the official data in use on the workflow server at EPA headquarters. The replication process is also unidirectional and not continuous, such that "hacking" into the database of record is not possible. This is illustrated in Figure 4-2.

Figure 4-1. Public Access to Health and Safety Studies

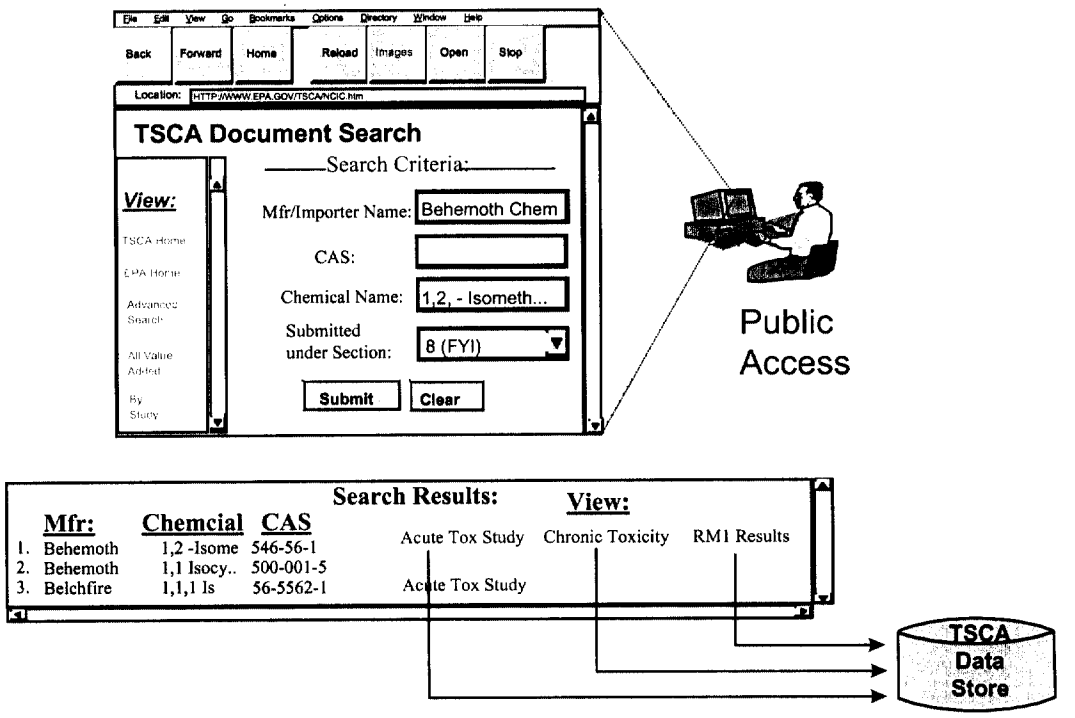
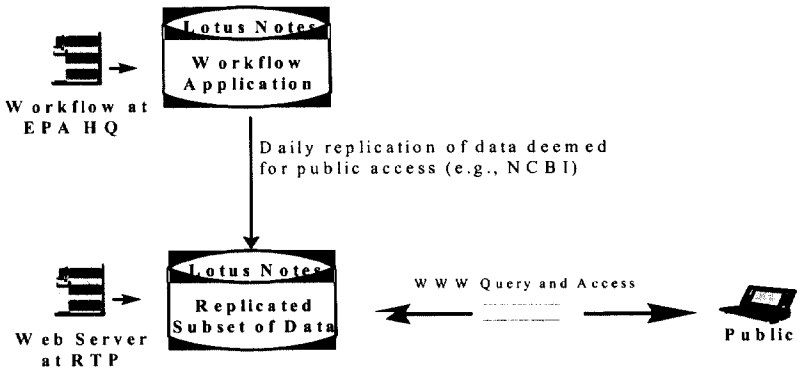


Figure 4-2. Basic Public Access Architecture



If desired, the index data could still be downloaded to magnetic tape and the health and safety submissions converted to microfiche for government and public use of the data. These sources include the National Library of Medicine, the National Technical Information Service, Chemical Information Systems, Inc., and the National Institute for Occupational Safety and Health. The RDMB retains all paper submissions and microfiche on file in the public docket.

PMN

Making sanitized PMN data available to the public is simply part of the workflow process. The most logical approach is to emphasize distribution over the web, similar to the Health and Safety process.

Chapter 5
Migration Steps and Issues

The business processes described in the previous chapters represent a jump to the "to-be" process that OPPT should be capable of achieving in less than three years time. The path to the to-be state may be incremental, however, and there are issues that OPPT must address to continue processing TSCA data in the interim.

Data Security

The biggest impediments to implementing an electronic architecture may be related to data security, some of which are only perceived, others of which are very real. Hardware and software security has progressed to the point where electronic commerce may be much more secure than paper processing. Real, significant issues remain, however, and these are complicated by the fact that enterprise-wide security is currently a very dynamic field. The issues discussed below are very basic and must be addressed, if not resolved, before electronic TSCA reporting can begin in earnest.

Central Registration

Any exchange of information with industry is going to require that OPPT authenticate the source of that data. At the very basic level in a web architecture this requires a user name and password. But how are these assigned and maintained? The answer to this can range from a low-level authentication process where a user can provide their identity as anything they choose and also select their own password to a sophisticated biometric system which verifies an individual's identity through something like a fingerprint. Whatever procedure is used, it is common sense that TSCA would not require a user to have a different password and identity for submitting data under different sections of TSCA.

In fact, the OEI is currently addressing an Agency-wide solution for a central registration process that gives an individual submitter one identity for all types of submissions to EPA. Because of the scope of this effort however, it is likely that OPPT can implement a solution for TSCA in a much shorter period of time with the idea of merging with the OEI solution when it becomes available. OPPT should establish a TSCA Central Registration infrastructure. This would consist of the following elements:

* A web page where potential submitters identify themselves requesting a TSCA user name and password. Data elements on this page would likely include:

◆ Name;

◆ Organization;

◆ Division or subunit within organization (down to Dun and Bradstreet number level)

◆ Address; and

◆ Phone/Fax/E-mail.

* A process, likely run by the Records & Dockets Management Branch, where the web requests are received and reviewed for rejection or approval. Because no CBI will initially be sent or accessed over the web, it would be sufficient to then send a user name and initial password to the person at that address via the web. If necessary, cursory checks could be made such as, for example, calling the person to verify phone numbers, or running the D&B number. That person would then access the web page and change the initial password to their own personal password.

* A database, on the back end of the web page, where identities are maintained. This would contain, for example, information about each submitter, when the passwords were due to expire, etc.

DIGITAL SIGNATURES

Because TSCA sections 4, 5, and 8 submissions require a signature, OPPT must wrestle with preserving non-repudiation in an electronic environment. Because of enforcement potential, this issue can impact far beyond OPPT and outside the Agency. Department of Justice, the Office of Management and Budget, and the Federal electronic commerce workgroups are all involved in this issue, and while there is a growing acceptance of digital signatures as a way to bind individuals to documents and transactions, lack of judicial case history has left little in the way of guidance for what is acceptable and what is not.

Fortunately the Federal Government has recently provided some guidance. The Government Paperwork Elimination Act (GPEA) of 1998 establishes the Office of Management and Budget the executive agent for the Federal Government's use of technology. In addition, GPEA defines procedures for use and acceptance of electronic signatures. The characteristics of an acceptable electronic signature, as defined by GPEA, include the use of accepted standards, technical neutrality, reliability of signature proportional to requirement, and the use of electronic acknowledgements. (Large filing programs, defined as more than 50,000 submissions, must also provide multiple methods for meeting the signature

requirement. TSCA is not a large filing program.)

The new functionality provided by Adobe Acrobat 4.0 meets these tests and meshes well with the advantages of storing PMNs and Health and Safety Studies in PDF. The new release of Acrobat 4.0 allows digital signatures to be placed on these documents. The digital signature functionality has the following capabilities:

* Handle multiple signatures;

* Accept a variety of third-party digital signature types;

* Date/Time stamps signatures and provides locations and reasons for signing;

* Detects any changes in document made after each signature and allows for "rollback" to see what was changed;

* Permits real-time validation of signature in public-key infrastructure (PKI) environment or local public key address book;

Because the signature capability in Adobe Acrobat 4.0 makes use of third party "plug-ins," this passes the technical neutrality test. In addition, the mechanics of the digital signature are specific to the third party solution, all of which generally conform to the standards used by these technologies. These include the X.509 standard for digital certificates, RSA or DSS algorithms for signing, and the SHA-1 and MD5 standards for secure hash algorithms.

CERTIFICATE MANAGEMENT

State of the art data security uses the public-private key approach for confidentiality (encryption) and non-repudiation (digital signatures). A user has a public key for encrypting and a private key for signing that only he knows. Others use that person's public key to send an encrypted message, for example. This arrangement works very well, provided you can be sure that you can match public keys and individuals (or organizations or servers) with confidence. In order to maintain that confidence, certificates are used which validate identities and public keys.

One approach that is being explored at the Federal level is the use of a Public Key Infrastructure (PKI). PKI is a way of establishing a "trusted" network of users. The administrator of the PKI issues certificates to all users in the network which validate their identity. Certificates are revoked when a user is no longer part of that network (e.g., leaves the organization). Several companies provide PKI services and software, including the two largest, Verisign and Entrust. Verisign issues and manages certificates as a service provider, while Entrust sells the

software necessary for an organization to run a PKI and issue their own certificates.

Instead of building a PKI, the best alternative may be for OPPT to take advantage of service providers such as Verisign. Since Adobe Acrobat 4.0 supports multiple service providers, it allows OPPT to stay out of the certificate management business. The government has standardized these services through the work and direction of the Federal PKI Steering Group and GSA's Access Certificates for Electronic Services (ACES) program. The ACES program maintains a schedule for acquiring certificate management services for various vendors including Verisign. Verisign describes the classes of digital ID's as follows:

* A Class 1 digital ID provides an individual with an unambiguous name and e-mail address. Generally it only verifies that the individual and e-mail address are bound.

* A Class 2 digital ID provides more identity assurance by third party verification of name, address, and other information.

* A Class 3 digital ID provides higher identity assurance because it requires that the requestor appear before a notary to have the application authenticated.

* A Class 4 digital ID is similar to the Class 3 digital ID, but it also verifies that the individual has a relationship with a specific organization.

To complete the analogy to paper documentation, the Class 1 digital ID may be like an electronic library card (in terms of robustness), a Class 2 may be akin to an electronic driver's license, and a Class 3 or 4 can be thought of as electronic passports. With this framework, OPPT would not need to set itself up as a certificate or registration authority. It could simply require that TSCA 4, 5, and 8 submissions be electronically signed with a class 3 or 4 digital ID chosen from a list of approved certificate authorities.

SECURE NETWORKING

In Chapter 3, we discuss the need for public access as well as the need to protect CBI. We propose the use of two Notes/Domino servers. The first server would be configured to allow industry to submit information in a secure environment (using SSL v3 as the underlying authentication protocol). The second server would be configured for public access and would therefore not require an ID and Password to gain access. However, content on the second server would be limited to NCBI only, which would have been obtained via periodic one-way replication (or a push) from the first server to the second server. In this fashion, CBI information

would never be open to public access.

Submission Tracking

As documents are received by OPPT, the Records and Dockets Management Branch (RDMB) must establish document tracking control. Until the Confidential Business Information Tracking System (CBITS) has been replaced by the to-be system, the to-be system must replicate the document tracking data required by CBITS. That data, consisting primarily of the document number and the document control number, must be entered along with the remaining data elements currently maintained in CBITS. The data need only be entered once for the to-be system and replicated for use by CBITS.

TSCA Data Store Implementation

The level of effort and coordination required to develop a core TSCA data store is sufficient that money and timing issues likely require a phased approach to construction. We describe one such approach in the following sections.

Phase 1: Development of an OPPT Enterprise Target Architectural Framework

The primary objective of this phase is to determine the appropriate target architecture for OPPT which will serve as the direction and decision framework for all OPPT information technology implementation decisions. To accomplish this, the following subtasks must be performed:

* Identify appropriate architectural principles as well as the business and technology drivers that affect them. Examples of sources of these principles are legislation, such as the Clinger-Cohen Act, EPA architectural guidance, the OPPT mission and others.

* Review all system requirements and related documentation developed by OPPT. This will include the OPPT standards for data development and display.

* Conduct an inventory of all of the potential databases and applications for inclusion in the consolidated database. This is inclusive of existing and planned Lotus Notes workflow applications as well as local desktop databases that may be maintained by OPPT members. Apeendix A identifies an OPPT list of databases that currently exist.

* Conduct analysis of the inventories and develop recommended target architecture.

* Prepare a detailed systems design master document as a roadmap to the end state of a consolidated data store. This document presents the target architecture and reflects the architectural principles that support a decision-making framework. As follow on phases are implemented the documentation should be updated accordingly with physical and normalized models replacing conceptual models.

* Conduct a Risk analysis of the recommended target architecture. This will serve as the baseline for risk assessments throughout the project.

Phase 2: Iteration 1 of Database Consolidation

In this iteration, the decision framework and architectural guidance from Phase 1 should be implemented. The goal of this phase is to incrementally field a capability for OPPT to achieve a consolidated data store. In this phase the groundwork for follow on iterations is established through the following:

* Project Management: The project plan, data management plan, and configuration management plan should be developed as the guideline for overall management during this and all following phases.

* Analysis/Design: The TSCATS 2.0 Lotus Notes database should be used to develop a conceptual and physical data model that will serve as a working prototype relational database using the Oracle Relational Database Management System (RDBMS). The documentation artifacts from this prototype serve as the baseline for future iterations and changes or enhancements. This document should include the entity relationship diagram (ERD), conceptual data model, physical data model and data dictionary. Prior to development, test scenarios should be submitted which address the approved requirements and constitute the majority of acceptance criteria for OPPT

* Development: The capability of the legacy applications that were used to populate the prototype database should be modified, enhanced or rebuilt. A reporting and ad hoc query capability should be developed to compliment the application capabilities and meet OPPT's requirements.

* Testing: Test scenarios should be implemented to ensure functional requirements have been met. Testing applications should be employed to simulate load, traffic and bandwidth metrics.

* Implementation: OPPT should install the iteration 1 incremental capability on production servers. End user and systems administration documentation and training should be provided. The master design document, the configuration management plan, the data management plan and the overall project plans should be revised and updated.

Phase 3: Iteration 2 of Database Consolidation

In this iteration, additional data from selected applications should be added to the consolidated data store. These added data elements are anticipated to originate from the Chemical Update System (CUS), Confidential Business Information Tracking System (CBITS), MEGA, PAIR, Management Information Tracking System (MITS), Section 5 New Chemicals (PENTA), CrTK, CORR, and all NCIC TSCA databases. Corresponding functionality from those applications should be added or mapped to the data store. For consistency the decision framework and architectural guidance from Phase 1 should be used. The goal of this phase should be to field the second incremental capability, building upon the delivered capability from Iteration 1, that will allow OPPT to achieve a consolidated data store. The same approach used to manage Iteration 1 of database consolidation should be used to manage Iteration 2 and update the documentation developed during Iteration 1.

Phase 4: Iteration 3 of Database Consolidation

In this iteration, additional data from selected applications should be added to the consolidated data store. These added data elements are anticipated to originate from the Export Notice tracking System (ENTS), Premanufacture Notice (PMN) database, and ISIS. Corresponding functionality from those applications should be added or mapped to the data store. For consistency the decision framework and architectural guidance from Phase 1 should be used. The goal of this phase should be to field the third incremental capability, building upon the delivered capability from Iteration 2, that will allow OPPT to achieve a consolidated data store. The same approach used to manage previous iterations of database consolidation should be used to manage Iteration 3 and update the documentation developed during Iteration 2.

Phase 5: Iteration Four of Consolidation; Communication Enhancements

In this iteration, EDI/UN/EDIFACT transaction set(s) or other appropriate techniques should be developed for the consolidated data store to insure seamless communication between U.S. and international entities using complementary systems. Transaction sets should be developed through coordination with appropriate stakeholders. For consistency the decision framework and

architectural guidance from Phase 1 should be used. This phase should be used to address any unresolved issues related to the consolidated data store and implement any enhancements approved by the CCB. The goal of this phase should be to field the fourth incremental capability, building upon the delivered capability from Iteration 3, that will allow OPPT to achieve a consolidated data store. The same approach used to manage previous iterations of database consolidation should be used to manage Iteration 4 and update the documentation developed during Iteration 3.

Chapter 6
Process Management

Identify Program Management and Assign Data Ownership

Before significant change is possible in TSCA processing, a change is also required in process management. The organization of OPPT necessitates cross-divisional processing which, in and of itself, is not a problem. A culture of decentralization, such as the EPA fosters, can be very productive particularly when individual analysis and innovation are required. The problem arises, however, when questions regarding overall program management occur. It is our assessment that the lack of clear responsibility and authority among and between programs hinders both current processing and the chances for process improvement.

Data ownership must also be clearly established. While program manager and data owner may be one in the same, the individual responsible for data quality and availability must be identified. With OPPT's role as TSCA information broker, the data owners become the primary customers within the office. Others in the office, including information management specialists, work toward satisfying the data owners.

Receiving Roles and Responsibilities

Roles and responsibilities for electronic receipt of TSCA submissions are discussed in the following sections. In general, OPPT must establish the electronic architecture for receipt (and processing) with hardware, software, and business rules. Responsibilities for these fall to the Information Technology and Support Branch, the Information Access Branch, and the Records and Dockets Management Branch, of the Information Management Division, respectively.

Information Technology and Support Branch

The Information Technology and Support Branch (ITSB) will provide and maintain the infrastructure necessary for the reengineered process. This includes the enterprise hardware, enterprise software, and enterprise training. Specifically, the ITSB must address the following components of the infrastructure:

- Scanning Equipment. In order to digitize incoming hard copies, ITSB must investigate current scanning equipment and provide the

RDMB with the hardware and software necessary for a medium volume scanning operation.

◆ Secure Network. The ITSB must develop alternatives for handling secure information including those submissions sent over the web. Two alternatives include:

◆ Integrating the CBILAN and NCLAN into one network where secure information is accessed only by authorized individuals. Security hardware and software have progressed from the point where separation of the data was necessary to a point now where security can be maintained without segregating the infrastructure.

◆ Maintain two separate networks. OPPT can preclude public access to CBI by maintaining a separate NCBI network as described in Chapter 3. Internal EPA access for authorized users would be provided through the single, but separate network containing both CBI and NCBI.

◆ Database Engines including Oracle. ITSB must provide the software for the TSCA data store that will permit efficient, effective access to all TSCA-related information. Workflow applications will run-off of this data.

INFORMATION ACCESS BRANCH

The Information Access Branch (IAB) is responsible for providing the applications that will receive, process, distribute, and archive the submissions. It is also responsible for the overall information access schema. Toward effective TSCA and PMN processing, IAB must take the following steps:

◆ Coordinate with RTP to maintain front-end web site for data submission and data replication to workflow server at EPA headquarters.

◆ Develop the overall processing architecture and assign roles and responsibilities.

◆ Develop the TSCA data store using the agency standard database engine (Oracle).

◆ Develop, or modify, Health and Safety Data (HaSD) and CRTK applications in Lotus Notes that meet the following minimum functionality:

◆ Discretely indexes HaSD cover sheet (meta data) and decision data

such as company name, chemical name, and study type.

◆ Permits "value-adding" of functional data such as hazard ranking and final disposition.

◆ Permits attachment of any number of supporting files from the submitter.

◆ Automatically notifies key personnel by e-mail when significant steps have or have not occurred in the CRTK and RM1/2 processes.

◆ Interfaces with backend database engine (Oracle) and/or document management software.

➢ Develop Web Access. IAB should take the steps necessary to make NCBI and public access information available via the World Wide Web. It must establish an architecture that connects OPPT with both the submitter and the public (for access), securely, through RTP. Periodically, perhaps daily, the RTP Lotus Notes application is replicated to the same application running on an OPPT Lotus Notes server located in Washington. The data are also passed through to the TSCA data store using some ODBC connectivity tool such as Lotus' Notes Pump. The data are now available for processing in OPPT. For public access, the process works in reverse. NCBI data are replicated periodically back to RTP where they are available for public access. In this manner, the public is not allowed access to the actual data record at OPPT.

RECORDS AND DOCKETS MANAGEMENT BRANCH (RDMB)

The Records and Dockets Management Branch of the Information Management Division is responsible for several functions which currently take place in the CBIC and NCIC. In the new process, both for electronic and paper submissions, these functions change somewhat. There are several scenarios that may take place with the new submission options.

◆ Before a submitter is allowed to submit data electronically, that submitter must first register with OPPT for a user name and password and provide, as necessary, their public signing key used for digital signatures. RDMB must administer the process where registrations requests are received, reviewed, maintained, and archived. They must also provide the submitter with a user name and initial password. In short, the RDMB manages the authorization control list (ACL) for access to the TSCA submissions page.

◆ Acquire expertise, such as a database administrator, necessary for

running an industrial strength database and application.

◆ If the submission and cover sheet data are submitted electronically over the web, there is no action required by the RDMB because data validation is performed automatically by the workflow application.

◆ If the entire TSCA transaction is received as a paper submission, the RDMB first checks the application to see if the study or cover sheet has already been submitted into the application. If not, the RDMB would then scan the document and convert it to an electronic image. If the submitter provided a cover sheet, the RDMB would then complete the HTML screen for the cover sheet data and post the submission to the web site in the same manner as the manufacturer above. If no cover sheet was provided, the RDMB would enter very basic indexing data (manufacturer name, date received, etc.) and route to the appropriate data owner. For Health and Safety submissions, this could be the Existing Chemical Branch (ECB). The ECB would provide greater cover sheet indexing (e.g., study type) before further processing. The RDMB then archives the paper submission. For PMN submissions, this could be the New Chemicals Notice Management Branch (NCNMB).

◆ If a submission arrives on diskette or CD-ROM, the RDMB must ensure the electronic document is virus-free and readable by the application (PDF reader). If rejected, the RDMB must notify the submitter. The RDMB would then upload the file to the application, complete the very basic indexing data, and route to the ECB (or NCNMB) for full indexing. The RDMB then archives the diskette or CD-ROM.

In order to reduce the number of permutations of possible submission scenarios with Health and Safety submissions and simplify the tracking process, OPPT should manage web submissions with cover sheet and study together. In other words, submitters should be required to submit cover sheet data over the web for every study submitted.

Upon successful upload of the submission, the workflow application will reflect current processing status at every phase of the process. Subsequent responsibilities of the RDMB, discussed in the next paragraphs, are the same whether the document was transmitted electronically by the manufacturer or scanned and uploaded by the RDMB.

The RDMB first performs completeness and page sequence checks of each submission. Tracking numbers to include the document number and document control number, should be assigned to the document and entered along with any

other required index data not included in the submission. These data may best be captured by a separate data entry input form specifically designed for the data currently maintained in CBITS. Prior to approving submissions for further processing, the RDMB must also indicate which submissions contain confidential business information (CBI) to prevent public access to that data.

After approval by the RDMB, the workflow application will notify all relevant parties of a submission requiring further processing by scientific and technical branches in the OPPT. Upon approval by RDMB, the submission itself will be stored, simultaneously accessible by all relevant parties, according to business rules established by the program manager, such as ECB.

Scientific Information Branch (SIB)

The Scientific Information Branch is responsible for maintenance of the core TSCA data store and applications. SIB will develop new views, reports, querrie, and workflow required for the system. Specifically SIB will be responsible for:

- Application maintenance. SIB will be responsible for trouble shooting any application specific problems that occur. Help desk inquires on application specific problems will be routed to SIB. Problems relating to data will be routed to the data owner.

- Client Applications including Lotus Notes. SIB must provide canned and ad hoc query capability, report writing software and document management software for internal EPA client use.

- Workflow management applications. SIB must develop workflow applications to track and process TSCA submissions efficiently and effectively, and capture all resulting value-added data.

Processing Roles and Responsibilities

Existing Chemical Assessment Branch (ECAB)

Once notified by the workflow application that a new submission has been received, the ECAB will view or download the submission in order to perform technical and scientific reviews. The ECAB will complete HTML screens designed for its role in recording and publishing the results of all submissions reviewed. Part of this effort includes providing further index information that may not have been submitted with the study. Once submitted, that data is a part of the health and safety data submission that can be viewed by the public.

High Production Volume Chemical Branch (HPVCB)

Once notified by the workflow application that a new submission has been

received, the HPVCB will view or download the submission in order to evaluate submissions and perform hazard ranking, initial screening, and additional data development. The HPVCB will complete data entry forms designed for its role in recording and publishing the data it is responsible for developing. Once submitted, that data is a part of the health and safety data submission that can be viewed by the public according to the business rules to be established.

CHEMICAL INFORMATION AND TESTING BRANCH (CITB)

The CITB monitors the workflow application in order to manage and report the status and progress on the processing of TSCA section 4 and 8d health and safety submissions. The CITB will complete HTML screens designed for its role in recording and publishing progress reports and final test data that are not already resident in the system. The workflow application can be programmed to produce various electronic or hard copy management reports for relevant parties on a periodic or as required basis.

OUTREACH BRANCH, ENVIRONMENTAL ASSISTANCE DIVISION

This branch is responsible for management of the 12b export notification process. Staff of the TSCA Assistance Information Service (generally known as the TSCA Hotline) are responsible for receiving electronic and hardcopy notifications from manufacturers, generating the letters notifying the receiving countries, and reporting on notifications for EPA's Office of Enforcement. The Hotline staff will download non-CBI notifications from the Internet to the Lotus Notes Export Notification Tracking System (ENTS). They will manually enter data into the system for hardcopy and all submissions claimed CBI. ENTS will automatically generate letters from the database notifying the receiving countries. Staff will initiate the daily upload of the updated non-CBI ENTS information on the Internet.

NEW CHEMICALS NOTICE MANAGEMENT BRANCH

The New Chemicals Notice Management Branch (NCNMB) is the first office notified by the RDMB through the workflow application that a PMN submission has been received. The NCNMB reviews the submission and all available supporting data and decides whether the new chemical should be processed through standard review, dropped or handled in some other manner. If the new chemical is to be processed through standard review, the NCNMB notifies the Industrial Chemistry Branch (ICB) through the workflow application that a new submission has been received. The NCNMB also performs the Risk Management/Regulatory control Function which includes developing the necessary documentation relative to how a new chemical is to be managed and regulated. The resulting documentation is attached to the PMN submission and can be viewed by the EPA internally and the public according to the business

rules to be established. The required documentation is also electronically sent by the workflow application to the PMN submitter.

INDUSTRIAL CHEMISTRY BRANCH

Once notified by the workflow application that a new submission has been received, the ICB will identify PMNs for review and schedule meetings for chemistry assessment. The ICB will complete HTML data entry forms designed for its role in establishing the plan for assessing the chemical, determining the need for additional reviews, and generating necessary reports to include the Initial Review Chemistry Report (IRCR). Once submitted, the chemical review scheduling data is stored in the Lotus Notes Workflow application that can be viewed by the relevant branches within OPPT. The report data is attached to the PMN submission and can be viewed by the EPA internally and the public according to the business rules to be established.

CHEMICAL ENGINEERING BRANCH

Once notified by the workflow application that the ICB has completed its review of a PMN submission, the Chemical Engineering Branch (CEB) reviews the PMN and the IRCR. The CEB will, as required, complete HTML data entry forms that simulate the Initial Review (IR) Engineering Report. Once submitted, that data is attached to the PMN submission and can be viewed by the EPA internally and the public according to the business rules to be established.

EXPOSURE ASSESSMENT BRANCH

Once notified by the workflow application that the CEB has completed the IRCR, the Exposure Assessment Branch (EAB) reviews the PMN, the IRCR, and the IR Engineering Report. The EAB will, as required, complete HTML data entry forms that simulate the IR Exposure and Fate Reports. Once submitted, that data is attached to the PMN submission and can be viewed by the EPA internally and the public according to the business rules to be established.

NEW CHEMICALS SCREENING AND ASSESSMENT BRANCH

Once notified by the workflow application that a new submission has been received, the New Chemicals Screening and Assessment Branch (NCSAB) reviews the PMN to determine the hazard level of the new chemical. The NCSAB will complete an HTML data entry form that documents the hazard level. Once notified that the EAB has completed the IR Exposure and Fate Reports, the NCSAB reviews that Report and completes HTML data entry forms that simulate the Structure Activity Relationships Team (SAT) Report. Once submitted, the hazard level and SAT are attached to the PMN submission and can be viewed by the EPA internally and the public according to the business rules to be